[redacted]

NCT : NCT02268253

STATISTICAL ANALYSIS PLAN

STUDY SPONSOR:	Stemline Therapeutics, Inc. 750 Lexington Avenue-11th Floor New York, NY 10022 Tel: 646-502-2310
SPONSOR REPRESENTATIVE:	[redacted]
PROTOCOL TITLE:	*Tagraxofusp (SL-401) in Patients with Chronic Myelomonocytic Leukemia (CMML)*
PROTOCOL NUMBER:	STML-401-0314
PROTOCOL VERSION AND DATE:	Amendment 9: 30 September 2021 Amendment 8: 05 May 2020 Amendment 7: 21 October 2019 Amendment 6: 25 June 2018 Amendment 5: 27 March 2018 Amendment 4: 10 June 2017 Amendment 3: 24 June 2016 Amendment 2: 08 July 2015 Amendment 1: 13 January 2015 Original: 24 September 2014
NAME OF TEST DRUG:	Tagraxofusp (SL-401)
PHASE:	Phase 2
METHODOLOGY:	Non-Randomized, Open-Label
ANALYSIS PLAN DATE:	30 November 2023
ANALYSIS PLAN VERSION:	Final Version 2.0
AUTHOR:	[redacted]

Confidentiality Statement

The information contained herein is confidential and the proprietary property of Stemline Therapeutics, Inc. and any unauthorized use or disclosure of such information without the prior written authorization of Stemline Therapeutics, Inc. is expressly prohibited.

APPROVAL SIGNATURE PAGE

Protocol Title: Tagraxofusp (SL-401) in Patients with Chronic Myelomonocytic Leukemia (CMML)

Sponsor: Stemline Therapeutics, Inc.
750 Lexington Avenue-11th Floor
New York, NY 10022

Protocol Number: STML-401-0314

Document Date / Version: 30 November 2023 / Final Version 2.0

Veristat, Inc. Author: [redacted]

Approval Signature	Job Title
[redacted]	[redacted]

Sponsor Signatory: [redacted]

Sponsor Approval

By signing this document, I acknowledge that I have read the document and approve of the planned statistical analyses described herein. I agree that the planned statistical analyses are appropriate for this study, are in accordance with the study objectives, and are consistent with the statistical methodology described in the protocol, clinical development plan, and all applicable regulatory guidances and guidelines.

I have discussed any questions I have regarding the contents of this document with the biostatistical author.

I also understand that any subsequent changes to the planned statistical analyses, as described herein, may have a regulatory impact and/or result in timeline adjustments. All changes to the planned analyses will be described in the clinical study report.

Approval Signature	Job Title
[redacted]	[redacted]

TABLE OF CONTENTS

Section **Page**

TABLES INCLUDED IN THE TEXT

Page

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviation	Definition
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine aminotransferase
AML	Acute myeloid leukemia
AST	Aspartate aminotransferase
ATC	Anatomic therapeutic class
AUClast	Area under the concentration time curve from time = 0 to the time of the last quantifiable concentration
AUCinf	Area under the concentration time curve from time = 0 extrapolated to infinity
AUMClast	Area under the first moment of the concentration-time curve from time =0 to the time of the last quantifiable concentration
BLQ	Below the limit of quantification
C	Cycle
CI	Confidence interval
CL	Total body clearance
CLS	Capillary leak syndrome
Cmax	Maximum observed concentration
CMML	Chronic Myelomonocytic Leukemia
CR	Complete response
CSR	Clinical study report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
D	Day
DLT	Dose-limiting toxicity
DOR	Duration of response
ECOG PS	Eastern Cooperative Oncology Group Performance Status
ECG	Electrocardiogram
eCRF	Electronic case report form
ELN	European Leukemia Net
hIL-3	Human interleukin-3
HMA	Hypomethylating agents
HU	Hydroxyurea
IL-3R	Interleukin-3 Receptor
IWG	International Working Group
IWG-MRT	International Working Group Myeloproliferative Neoplasms Research
IV	Intravenous
[redacted]	[redacted]

Abbreviation	Definition
λz	First order terminal elimination rate constant
MDS	Myelodysplastic syndrome
MedDRA	Medical Dictionary for Regulatory Activities
MF	Myelofibrosis
mITT	Modified intent-to-treat
MPN	Myeloproliferative neoplasm
MPN-SAF-TSS	Myeloproliferative Neoplasm Symptom Assessment Form Total Symptom Score
MRI	Magnetic resonance imaging
MRTinfIV	Mean residence time extrapolated to infinity
MTD	Maximum tolerated dose
MTeD	Maximum tested dose
NCA	Non-compart mental analysis
ORR	Objective response rate
OS	Overall survival
PCS	Potentially clinically significant
pDC	Plasmacytoid dendritic cell
PD	Progressive disease
PED	Primary Eosinophilic Disorders
PFS	Progression-free survival
PK	Pharmacokinetic
PP	Per-Protocol
PR	Partial response
PT	Preferred term
Rel Day	Relative study day
RP2D	Recommended phase 2 dose
SAE	Serious adverse event
SAP	Statistical analysis plan
SCT	Stem Cell Transplant
SD	Stable disease
SM	Systemic Mastocytosis
SMQ	Standardized Medical Query
SOC	System organ class
t½	Terminal elimination half-life
TEAE	Treatment-emergent adverse event
TI	Infusion duration
tlast	Time of last quantifiable concentration
tmax	Time of first occurrence of Cmax
[redacted]	[redacted]

Abbreviation	Definition
Vss	Steady-state volume of distribution
Vz	Volume of distribution during the terminal disposition phase
WHO	World Health Organization

1. INFORMATION FROM THE STUDY PROTOCOL

1.1. Introduction and Objectives

1.1.1. Introduction

Certain subtypes of myeloproliferative neoplasms (MPNs) and myelodysplastic syndrome (MDS)/MPNs represent clonal proliferations arising from malignant myeloid stem cells with propensity to overexpress interleukin-3 receptor (IL-3R), rendering them potentially susceptible to tagraxofusp. MPNs are associated with considerable morbidity and mortality, and risk of transformation to acute myeloid leukemia (AML). The substantial IL-3R expression on mast cells and eosinophils suggests that tagraxofusp may have the potential to diminish tumor bulk in addition to relevant underlying stem cell populations. The visible presence of plasmacytoid dendritic cell (pDC) nodular infiltrates in the bone marrow of a subset of chronic myelomonocytic leukemia (CMML) patients suggests a potential role of supporting dendritic cells within some of these neoplasia; preclinical data from other myeloid malignancies has indicated that targeting of these IL-3R expressing supporting cell populations may augment tagraxofusp antineoplastic effects observed against malignancies characterized by IL-3R expression (Chauhan 2014).

CD123 expression has been detected on CMML blasts, CMML monocytes, and leukemia stem cells in both MPNs and CMML (Shen 2015, Krishnan 2018). In one study of 20 CMML patient samples, CD123 was found to be expressed at higher levels on blasts and monocytes compared to lymphocytes, which showed low to no expression (Krishnan 2018). Moreover, CD123 expression on both blasts and monocytes was higher in CMML-1 and CMML-2 patients relative to CMML-0 patients (Krishnan 2018). Another study assessing 118 CMML bone marrow samples by a validated multi-color flow cytometry immunophenotyping assay found that CD34+ blasts exhibited increased expression of CD123 in 64% of CMML patients compared to 3% of control patients (Shen 2015).

CD123 is also expressed on neoplastic pDCs in the tumor microenvironment of CMML (Facchetti 2016, Lucas 2019, Vermi 2004, Brunetti 2017, Ji 2014, Naresh 2010, Vermi 2011). These pDCs have been shown to share mutations with, and belong to, the malignant CMML clone (Lucas 2019). In particular, pDCs have been shown to share RAS pathway mutations with the CMML clone, which are associated with poor prognosis including a significant increase in the cumulative risk of AML transformation (Lucas 2019). Notably, RAS pathway mutations are associated with a higher fraction of pDCs in the bone marrow (> 5%) of patients with CMML and clinically manifest as proliferative disease (Lucas 2019, Ricci 2010). Consistent with these findings, pDCs have been identified in the spleen of some patients with CMML, and it has been suggested that pDCs could serve as a therapeutic target in patients with CMML (Lucas 2019, Pophali 2018).

1.1.2. Study Design

Study STML-401-0314 was initially designed as a 2-stage, non-randomized, open-label, multicenter study of tagraxofusp in patients with subtypes of advanced, high risk MPNs. Initially in this study, patients with CMML and myelofibrosis (MF), as well as systemic

mastocytosis (SM) and advanced symptomatic chronic eosinophilic leukemia, including primary eosinophilic disorders (PED) were eligible for enrollment.

In Stage 1, which is complete at the time of this Statistical Analysis Plan (SAP), only patients with CMML and MF were enrolled. The primary objective of Stage 1 was to establish the maximum tolerated dose (MTD) or the maximum tested dose (MTeD) at which multiple dose-limiting toxicities (DLTs) were not observed. In Stage 2, further characterization of the safety and initial assessment of clinical activity at the recommended phase 2 dose (RP2D) were intended in patients with MF and CMML, with an initial planned sample size of 18 patients in each cohort, then extended to approximately 30 patients in each cohort. Enrollment in Stage 2 is closed.

In Stage 1, 9 patients (5 with CMML and 4 with MF) received tagraxofusp at doses of [redacted] via intravenous (IV) infusion on days 1-3 every 21 days for Cycles 1-4, every 28 days for Cycles 5-7, and every 42 days for Cycles 8 and beyond. No MTD was identified in Stage 1 and the MTeD was 12 μg/kg which was considered the RP2D. In Stage 2 of this study, patients received tagraxofusp at the RP2D, 12 μg/kg, on the same schedule as Stage 1. Stage 2 primarily included patients with CMML and MF; 1 SM patient was treated during Stage 2 at the RP2D.

Based on preliminary clinical activity observed and increasing scientific and clinical knowledge of CMML, Stage 3A was added by protocol amendment 7, dated October 21, 2019. Stage 3A enrolled 2 populations of patients with CMML-1 and CMML-2:

1. Previously untreated patients who are high risk and not expected to benefit from hypomethylating agents (HMAs) (previously untreated patients; approximately 20 patients), and

2. Patients who are primary or secondary failures to first-line therapy, including HMAs, hydroxyurea (HU), or intensive chemotherapy (first-line failure patients; approximately 20 patients).

1.1.3. Study Objectives

This SAP is designed to outline the methods to be used in the analysis of study data in order to answer the study objective(s). Populations for analysis, data handling rules, statistical methods, and formats for data presentation are provided. The statistical analyses and summary tabulations described in this SAP will provide the basis for the results sections of the clinical study report (CSR) for this trial.

This SAP will also outline any differences in the currently planned analytical objectives relative to those planned in the study protocol.

1.1.3.1. Stage 1 Study Objectives

The primary objectives of Stage 1 of the study were to determine the MTD or the MeTD at which multiple DLTs are not observed of tagraxofusp and to characterize the safety profile of tagraxofusp. Stage 1 objectives are described in Protocol Amendments 6 and earlier.

1.1.3.2. Stage 2 Study Objectives

The secondary objectives for the study from Protocol Amendment 6 and earlier are to characterize the anti-tumor activity (including response rates, components of the response criteria, duration of response [DOR], progression-free survival [PFS], and overall survival [OS]), pharmacokinetic (PK), and immunogenicity of tagraxofusp.

An additional primary objective of Stage 2 is a preliminary assessment of response rates associated with SL-401 in each of the evaluated MPNs.

Specific components of the International Working Group Myeloproliferative Neoplasms Research and Treatment (IWG-MRT) consensus response criteria and additional clinical parameters for CMML will be assessed as secondary endpoints, including:

- Change in body weight (individual and median) over time.
- Change in hemoglobin (individual and median) over time.
- Change in Eastern Cooperative Oncology Group Performance Status (ECOG PS) (individual and median) over time.
- PFS (as defined in IWG-MRT consensus).
- OS.

Specific components of the IWG-MRT/ European LeukemiaNet consensus response criteria and additional clinical parameters for MF will be assessed as secondary endpoints, including but not limited to:

- Proportion of patients with reduction of ≥35% in spleen volume at week 24 (postbaseline, as measured by magnetic resonance imaging [MRI] or computed tomography [CT]).
- Proportion of patients with reduction of ≥35% in spleen volume at week 48 (postbaseline, as measured by MRI or CT).
- Proportion of patients (transfusion-dependent at baseline) who achieve transfusion independence (during any 12-week period following initiation of study therapy, as defined in IWG-MRT consensus).
- Proportion of patients who report a reduction in total symptom score of ≥50% from baseline to week 24 (as assessed via the Myeloproliferative Neoplasm Symptom Assessment Form Total Symptom Score [MPN-SAF TSS]).
- Change in hemoglobin (individual and median) over time.
- Change in body weight (individual and median) over time.
- Change in ECOG PS (individual and median) over time.
- PFS (as defined in IWG-MRT/ELN consensus).

- OS.

Exploratory objectives are to characterize expression of IL-3R/CD123 (and other potentially relevant stem cell and disease markers) on myeloid malignant cells and associated cell populations in the bone marrow (BM) (when feasible), to evaluate potential changes in IL-3R/CD123 (and other potentially relevant marker) expressing populations over time, and preliminary correlation of baseline IL-3R/CD123 (and other potentially relevant markers) expression and clinical efficacy (including response rates).

Note that from Amendment 7 and later of the protocol, which focused on Stages 2 and 3a of the study, the objectives were not specific with regard to primary versus secondary objectives.

Objectives of Stage 2 in patients with CMML from Protocol Amendment 7 and later were:

- To characterize the safety profile of tagraxofusp.
- To obtain preliminary evidence of response to tagraxofusp therapy as assessed by the International Working Group (IWG) MDS consensus response criteria (Cheson 2006), and to obtain data relative to additional clinical parameters that may be used in determination of therapeutic benefit.

Objectives of Stage 2 in patients with MF from Protocol Amendment 7 and later were:

- To characterize the safety profile of tagraxofusp.
- To evaluate the potential efficacy of tagraxofusp based on evaluation of specific components of the IWG-MRT and European LeukemiaNet consensus response criteria (Tefferi 2013B) and additional clinical parameters.
- To collect clinical data and evaluate the activity of tagraxofusp in specific sub-populations of patients with MF that are of interest, including patients who have relapsed from or are refractory to treatment with a JAK inhibitor (JAKi), patients who are not candidates for treatment with JAKi (due to thrombocytopenia) and patients with concurrent monocytosis.

1.1.3.3. Stage 3A Study Objectives

Objectives of Stage 3A in patients with CMML are:

- To characterize the safety profile of tagraxofusp.
- To obtain preliminary evidence of response to tagraxofusp therapy as assessed by the MDS/MPN 2015 criteria (Savona 2015), with the addition of a stable disease (SD) category, and to obtain data relative to additional clinical parameters that may be used in determination of therapeutic benefit.

The MDS/MPN 2015 criteria (Savona 2015) represent the current consensus criteria for evaluation of clinical benefit in patients with CMML, beyond those specified in the IWG-MDS criteria (Cheson 2006). Thus, as an inherent component of the design of Stage 3A, responses to treatment are being assessed using the MDS/MPN 2015 criteria. The hematology community has recognized that a lack of a SD category in the 2015 MDS/MPN IWG criteria creates

considerable uncertainty in the categorization of patients on any CMML study. Therefore, a SD category, defined as those patients not belonging to any of the defined categories in the 2015 criteria, has been added as a response category.

The purpose of Stage 3A, is to analytically evaluate the clinical benefit derived from individual components, or variations thereof, of the MDS/MPN 2015 criteria that may allow for a prospective definition of a primary efficacy endpoint. Clinical benefit includes erythroid, platelet, neutrophil, spleen, and symptom responses, as defined in Table 8-4 in Section 8.2, as well as transfusion dependence status and symptom score.

1.2. Study Design

1.2.1. Synopsis of Study Design

This is a non-randomized, open-label, multicenter study, divided into multiple stages.

1.2.1.1. Stage 1 Study Design

During Stage 1, approximately 12-36 patients were to be treated with tagraxofusp at doses of [redacted] /day for 3 consecutive days at the beginning of each 21-day cycle. Three to 6 patients were to be treated at each dose level. All patients within a cohort must have completed the first cycle of therapy before patients from a new cohort received tagraxofusp at the next higher dose. No intra-patient dose escalation was allowed. The first cohort of patients was to receive tagraxofusp at a dose of [redacted]. After all patients in this cohort completed the first cycle of therapy, the dose for the second cohort of patients was to increase by [redacted] to [redacted] /day, conditional on the DLT rules. Beginning with the [redacted] /day dose level of tagraxofusp, a decision to allow treatment at the next higher dose level would depend on the number of patients who experienced a DLT during the first cycle, as per the protocol. Expansion of the [redacted] /day cohort to up to 6 patients, if necessary, would follow the same rules as the [redacted] /day cohort. The same DLT rules also applied to the [redacted] /day, [redacted] /day, [redacted] /day, and [redacted] /day cohorts.

1.2.1.2. Stage 2 Study Design

Stage 2 was a non-randomized, open-label multicenter study stage to determine the safety and efficacy of single-agent tagraxofusp in patients with CMML and separately in patients with MF. Tagraxofusp was administered as a 15-minute IV infusion once daily on Days (D) 1 to 3 of a 21-day cycle (Cycles [C] 1 to 4); and on D1 to 3 of a 28-day cycle (C5 and beyond). The dosing period may have been extended for dose delays up to D10 of each cycle. Patients could continue to receive treatment with tagraxofusp as long as, in the opinion of the Investigator, benefit from treatment was indicated.

Beginning with Protocol Amendment 7, up to approximately 20-25 additional patients with MF were to be enrolled in Stage 2 to further elucidate the safety and efficacy profile of tagraxofusp in this patient population. Thus, across all study stages, approximately 50-55 patients with MF could be enrolled.

Note: CMML and MF patients consented and enrolled with Amendment 6 or prior, and currently on active treatment or in survival follow-up, continued to be followed according to the

schedule of events in Amendment 6. These patients were not reconsented under Amendment 7 or beyond.

1.2.1.3. Stage 3A Study Design

Stage 3A is a 2-arm, non-randomized, open-label multicenter study stage in patients with CMML intended to identify clinically important responses that connote clinical benefit and potential endpoints, as well as gather data on CD123 expression on pDCs and monocytes in CMML.

Stage 3A includes 2 populations of patients with CMML: those with CMML-1 or CMML-2 who are refractory/resistant/intolerant to HMAs, or HU, or intensive chemotherapy; and patients with treatment naïve CMML-1 or CMML-2 with molecular features associated with a poor prognosis who are not expected to benefit from HMAs.

In Stage 3A patients were administered tagraxofusp at a dose of [redacted] as a 15-minute IV infusion once daily on D1 to 3 of a 21-day cycle (C1 to C4) and then on D1 to 3 of a 28-day cycle (C5 and beyond). The dosing period may be extended for dose delays up to D10 of each cycle.

1.2.2. Randomization Methodology

As this is a single-agent study, randomization is not applicable.

1.2.3. Stopping Rules and Unblinding

Tagraxofusp treatment may be discontinued for any of the following reasons:

- Patient withdrawal of consent.
- Occurrence of unacceptable toxicity, including DLT.
- Tagraxofusp-related anaphylaxis or Grade ≥ 3 hypersensitivity reaction.
- Disease recurrence/progression.
- Intercurrent illness that prevents further administration of tagraxofusp.
- Patient non-compliance.
- Occurrence of pregnancy.
- Investigator's decision.

The reason for tagraxofusp discontinuation and the date of discontinuation was to be recorded in the electronic case report form (eCRF).

1.2.4. Study Procedures

The schedule of assessments, as outlined in the study protocol, is provided in Section 7, Table 7-1, Table 7-2, Table 7-3, and Table 7-4.

1.2.5. Efficacy, Pharmacokinetic, and Safety Parameters

1.2.5.1. Efficacy Parameters and Endpoints

Efficacy data from Stages 1 and 2 among patients treated at 12 ug/kg/day will be analyzed with descriptive statistical methods. Further details on the definitions and analysis methods for efficacy endpoints are provided in Section 4.3. Refer to Section 8 for detailed response criteria established for CMML and MF.

Primary Efficacy Endpoint:

The primary efficacy endpoint is investigator assessed objective response rate (ORR) for both CMML and MF. Objective response rate is defined as the number and percentage of patients who achieve disease-specific complete response (CR) or partial response (PR) after treatment.

Time to objective response and bridge to stem cell transplant after objective response will be assessed to support the primary efficacy endpoint.

Specific components of the IWG-MRT/ELN response criteria for defining clinical improvement for MF patients include anemia, spleen, and symptom responses. Specific components of the IWG-MRT response criteria for clinical benefit for CMML patients include erythroid, platelet, neutrophil, spleen, and symptom responses. The datapoints corresponding to each individual component of the responses will also separately be assessed and compared to baseline results for both MF and CMML patients as secondary endpoints. These include spleen volume, transfusion dependency, tumor symptoms score, blast results, absolute monocyte count results, white blood cell count results, hemoglobin results, and platelet results.

MF patients will be classified as “dependent” or “independent” at baseline for packed red blood cells (pRBC) if they received transfusions of at least 6 units in the 12 weeks prior to the study enrollment, for a hemoglobin level of < 85 g/L, in the absence of bleeding or treatment-induced anemia. In addition,the most recent transfusion episode must have occurred in the 28 days prior to study enrollment. Transfusion independence for MF will be defined as patients who experience at least a 12 week period during any point on treatment with no transfusion of that type required, including hemoglobin level of >=85 g/L.

CMML patients will be classified as “dependent” or “independent” at baseline within each category of platelets, packed red blood cells (pRBC), and both (i.e., receiving both platelets and pRBC) if they received transfusions of at least 4 units in the 8 weeks prior to the study enrollment. Transfusion independence for CMML will be defined as patients who experience at least a 8 week period during any point on treatment with no transfusion of that type required.

Secondary Efficacy Endpoints

The secondary efficacy endpoints are DOR, PFS, OS, and duration of follow-up.

Duration of objective response is defined as the time from when criteria are first met for CR or PR until the date that the criteria for relapse is met. Progression-free survival is defined as the time from the date of first infusion of tagraxofusp to the date of PD or death from any cause. Overall survival is defined as the time from the date of first infusion of tagraxofusp to the date of death from any cause.

1.2.5.2. Pharmacokinetic and Immunogenicity

Pharmacokinetics

[redacted] A quality control check of the NCA will be conducted.

Per kg dose levels will be converted to actual doses using subject body weight. Actual administered doses will be used in the calculation of PK parameters.

If pre-dose PK concentrations on Day 3 will result not-measurable in most subjects and very low in the few subjects with measurable PK concentrations, then single-dose PK parameters will be estimated for each PK profile (Cycle 1 Day 1 (C1D1), Cycle 1 Day 3 (C1D3), Cycle 4 Day 1 (C4D1), and Cycle 4 Day 3 (C4D3)).

The following PK parameters will be calculated, as data permit:

Name	**Description**	**Formula**
C_{max}	maximum observed concentration	
t_{max}	time of first occurence of C_{max}	
t_{last}	time of last quantifiable concentration	
AUC_{last}	area under the concentration time curve from time = 0 to the last quantifiable observation	Calculated with the *Linear Up/Log Down* method: Linear trapezoidal method $AUC_{(t1-t2)} = {(t2-t1)\cdot(C1+C2)}/{2}$ Logarithmic trapezoidal method $AUC_{(t1-t2)} = {(t2-t1)\cdot(C2-C1)}/{log_e(C2/C1)}$
λ_z	first order terminal elimination rate constant	Estimated by linear regression of logarithmically transformed concentration versus time data

Name	Description	Formula
$t_{½}$	terminal elimination half-life	$log_e(2)/\lambda_z$
AUC_{inf}	area under the concentration time curve from time = 0 extrapolated to infinity	$AUC_{last} + C_{last}/\lambda_z$ where Clast is the last observed quantifiable concentration
$\%AUC_{extrap}$	Percentage of AUC_{inf} obtained by extrapolation	$(AUC_{inf} - AUC_{last})/AUC_{inf} \cdot 100$
CL	total body clearance	$Dose/AUC_{inf}$
$MRT_{inf\,IV}$	mean residence time extrapolated to infinity	$(AUMC_{last} + t_{last}(AUC_{inf} - AUC_{last}))/AUC_{last} - TI/2$ where TI is the infusion duration, and AUMClast is the area under the first moment of the concentration-time curve from time zero to the time of the last quantifiable concentration.
V_{ss}	steady state volume of distribution	$CL_{ss} \cdot MRT_{\text{inf}\,IV}$
V_z	volume of distribution during the disposition phase	CL/λ_z

The following rules will be used for λ_z calculation:

(i) at least 3 non-zero time-points in the elimination phase for linear regression, excluding C_{max},
(ii) adjusted R squared ≥ 0.8,
(iii) span ≥ 2.

When criteria (i) or (ii) for λ_z calculation will not be satisfied, both λ_z and related parameters (i.e. $t_{½}$, AUC_{inf}, V_z, CL, $MRT_{inf\,IV}$ and V_{ss}) will not be calculated. Cases where span < 2 will be flagged.

If AUC_{extr} > 20 %, AUC_{inf} and related parameters (i.e. $MRT_{inf\,IV}$ and V_{ss}) will not be reported.

AUC_{last} will be calculated and reported for subjects with at least 2 non-zero/BLQ time-points.

Missing concentration will be left missing for the calculation of summary concentration statistics and for the estimation of PK parameters.

All concentration results reported as below the limit of quantification (BLQ) will be set to zero.

Erroneous and physiologically implausible individual PK concentration values might be excluded from the estimation of PK parameters and from the calculation of summary concentration statistics. Excluded data will be listed with the reason for exclusion provided and will be flagged in the individual concentration listings.

Actual times from the start of the Tagraxofusp infusion and actual infusion durations will be used in the estimation of PK parameters. For records where the sample collection time is missing, the scheduled collection time will be used. For records where the calculated actual time from the start of Tagraxofusp infusion will result in a negative value, the actual time will be set to zero. For records where the actual infusion duration is missing, the scheduled infusion duration of 15 min will be used. For ease of data presentation, scheduled collection times will be used in tables and figures of PK concentration data.

PK concentrations collected in unscheduled visits will be excluded from the analysis.

Immunogenicity

Summary of immunogenicity parameters, including anti-drug antibodies (further characterized for specificity to Diphtheria Toxin and/or human interleukin-3 [hIL-3]), specific anti-hIL-3 antibodies, and tagraxofusp neutralizing antibodies will be performed. Further details on these analyses are provided in Section 4.5.

1.2.5.3. Safety Parameters

Safety evaluations performed during the study included physical examinations, measurement of vital signs, 12-lead electrocardiograms (ECGs), clinical laboratory evaluations including hematology, serum chemistry, coagulation, and urinalysis, and monitoring of adverse events (AEs) (including DLTs and treatment discontinuations due to toxicity), and concomitant medications.

Further details on the definitions and analysis methods for safety endpoints are provided in Section 4.4.

2. PATIENT POPULATION

2.1. Population Definitions

The following patient populations will be evaluated and used for presentation and analysis of the data for all trial stages:

- Modified Intent-to-Treat (mITT) Population: All patients who are eligible based on the screening criteria who received at least 1 dose of tagraxofusp and are evaluable for efficacy (has an efficacy assessment after first treatment or discontinued due to death or progressive disease) prior to that time. Patients will be grouped according to the planned dose level at time of enrollment.

- Per-Protocol (PP) Population: All patients from the mITT population who are compliant with all major aspects of the protocol and have received at least 1 cycle of tagraxofusp (or treatment was delayed or patient discontinued treatment early due to safety finding related to study treatment). Patients will be grouped according to the planned dose level at time of enrollment.

- Immunogenicity Population: All patients enrolled in the study who received at least 1 dose of tagraxofusp and had at least 1 non-missing immunogenicity sample result.

- PK Population: All subjects who receive at least 1 dose of Tagraxofusp, have at least one plasma sample collected, and the corresponding plasma PK concentration is quantifiable.

- Safety Population: All patients enrolled in the study who received at least 1 dose of tagraxofusp. Patients will be grouped according to the actual dose level received.

The mITT population is the primary population for the analysis of efficacy parameters, and will be used to support baseline endpoints. The immunogenicity population will be used for summary of immunogenicity results. The PK population will be used for the analysis and summary of PK data. The Safety population is the primary population for the analysis of disposition, baseline, and safety endpoints. Patients included in the PP population or reasons for exclusion from the PP population will be identified in a data listing. Note that the intent-to-treat population, defined as all eligible patients, will not be used in any analyses.

2.2. Protocol Violations

At the discretion of the Sponsor, major protocol violations, as determined by a review of the data prior to database lock and the conduct of statistical analyses, may result in the removal of a patient's data from the PP population. The Sponsor or designee will be responsible for producing the final protocol violation file (formatted as a Microsoft Excel file), in collaboration with Veristat and the data monitoring group as applicable; this file will include a description of the protocol violation and clearly identify whether or not this violation warrants exclusion from the PP population. This file will be finalized prior to database lock for any interim and final analyses.

All protocol violations will be presented in a data listing.

3. GENERAL STATISTICAL METHODS

3.1. Sample Size Justification

3.1.1. Stage 2 Sample Size Justification

An objective of Stage 2 is a preliminary assessment of response rates associated with tagraxofusp in CMML and MF. Because available anti-cancer agents are associated with very limited responses in the MPNs evaluated in this study, it is anticipated that an agent that confers a response rate ≥ 10% may be considered active and worthy of additional investigation.

Initially, a sample size of 18 patients with a particular MPN (CMML or MF) was planned for enrollment in Stage 2. Furthermore, it was planned that in settings where there is evidence that a sufficient proportion of patients have experienced improvement in components of the response criteria, that additional investigation of tagraxofusp in a given MPN (CMML or MF) may have been considered appropriate based on preliminary data.

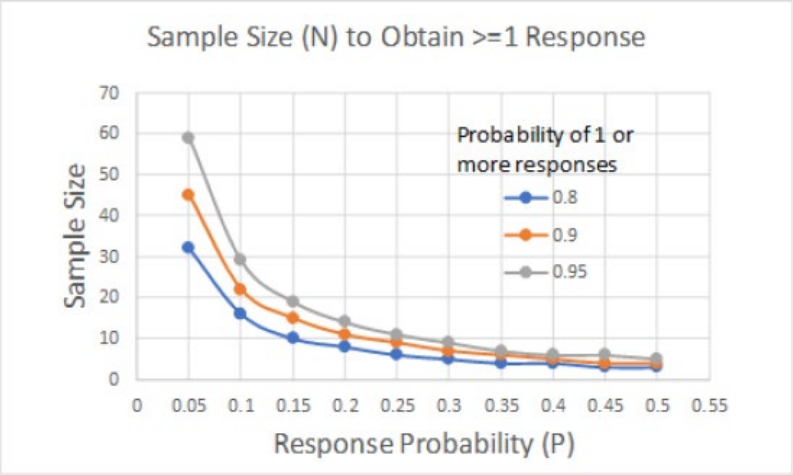
In order to ensure continued access to tagraxofusp in a controlled clinical study setting while additional investigations are being designed, the sample size in Stage 2 was increased by approximately 20-25 additional patients with MF in order to allow investigation of tagraxofusp in patients with MF and monocytosis and/or other emerging biological features that could help characterize the disease. Thus, across all study stages, approximately 50-55 patients with MF were planned to be enrolled.

3.1.2. Stage 3A Sample Size Justification

Approximately 20 patients with CMML may be enrolled in each of the two cohorts in Stage 3A to ensure a sufficient number of patients with specific baseline disease characteristics are enrolled to reject a meaningless rate of clinical activity in subsets of interest. (For our purposes, a meaningless rate will be taken to mean that there are no responders, using the definitions below, in the sample for each of the following subsets.) No formal stratification will be performed; however approximately 8-10 patients with baseline splenomegaly (defined as ≥ 5 cm below the left costal margin) and approximately 8-10 patients who are transfusion-dependent at baseline (receipt of at least 4 transfusions within 8 weeks before the first tagraxofusp dose) will be included in each 20-patient cohort. This intended enrollment distribution may be adjusted, depending on the actual prevalence of splenomegaly and transfusion-dependence seen in the general patient population as the study continues to enroll.

In this exploratory stage, the primary evidence of activity in each subset would be (1) a reduction in spleen volume that correlates with an improvement in symptom score or (2) conversion from transfusion-dependent to transfusion-independent. These occurrences on a per-patient basis would be considered responses within each subgroup. Clinical activity would have been demonstrated if at least 1 patient achieves either of these clinically important outcome responses. Up to 5 additional patients with either of these specific baseline characteristics may be enrolled in Stage 3A, to increase the confidence in the observed results, if necessary. It can be seen in the graph below that for any true rate of response that is ≥ 0.15, this approach would have at least 80% power to detect at least one responder. Given at least a single responder, additional methods such as confidence interval estimation or formal power calculations, may be used.

The operating characteristics of this approach to sample size estimation in Stage 3A may be seen in the following graph. This graph presents the sample size required to have either an 80%, 90%, or 95% probability of at least 1 responder when the sample size is N, and the true probability of response is P.



As an example, if the true response probability were 0.15, 10 patients evaluated for a given response type would result in at least one response with 80% probability.

3.2. General Methods

All data listings that contain an evaluation date will contain a relative study day (Rel Day). Pre-treatment and on-treatment study days are numbered relative to the day of the first dose of study medication which is designated as Rel Day 1. The preceding day is Rel Day -1, the day before that is Rel Day -2, etc. The last day of study medication is designated with an "L" (e.g., Rel Day 14L). Post-treatment study days are numbered relative to the last dose and are designated as Rel Day 1P, Rel Day 2P, etc.

All output will be incorporated into Microsoft Rich Text File or Excel files, or Adobe Acrobat PDF files, sorted and labeled according to the International Conference on Harmonisation (ICH) recommendations, and formatted to the appropriate page size(s).

Tabulations will be produced for appropriate demographic, baseline, efficacy, safety, and immunogenicity parameters. For categorical variables, summary tabulations of the number and percentage of patients within each category (with a category for missing data) of the parameter will be presented. Two-sided 95% confidence intervals (CIs) will be computed using the Clopper Exact method. For continuous variables, the number of patients, mean, median, standard deviation, minimum, and maximum values will be presented. Time-to-event data will be summarized using Kaplan-Meier methodology using 25th, 50th (median), and 75th percentiles with associated 2-sided 95% CIs, as well as percentage of censored observations.

3.3. Computing Environment

All descriptive statistical analyses will be performed using SAS statistical software Version 9.4, unless otherwise noted. Medical history and adverse events (AEs) will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 19.0. Concomitant medications will be coded using the World Health Organization (WHO) Drug Dictionary (September 2016).

3.4. Baseline Definitions

For all analyses, baseline will be defined as the most recent measurement prior to the first administration of study drug.

3.5. Methods of Pooling Data

Disposition, baseline, and safety summaries will be displayed by dose as well as by disease and trial stage. Tabulations presented by dose will include a summary of each dose level for MF and for CMML separately.

7 µg/kg/day	9 µg/kg/day	12 µg/kg/day	Overall

Tabulations presented by disease and trial stage will include a summary of each stage for MF (Stages 1 and 2) and for CMML (Stages 1, 2, and 3A).

MF			CMML			
Stage 1 (N=xx)	Stage 2 (N=xx)	Overall (N=xx)	Stage 1 (N=xx)	Stage 2 (N=xx)	Stage 3A (N=xx)	Overall (N=xx)

For myelofibrosis patients, summaries will be displayed by Dynamic International Prognostic Scoring System (DIPSS)-Plus Risk for treatment naïve and for relapsed/refractory patients separately.

Intermediate-1	Intermediate-2	High	Overall

For chronic myelomonocytic leukemia patients, summaries will be displayed by CMML Type for treatment naïve and for relapsed/refractory patients separately.

CMML-1	CMML-2	Overall

One patient with SM was treated in Stage 2 and will be presented in data listings. No patients with PED were treated so no separate tabulations or listings will be presented.

Summaries of efficacy results will be displayed separately for MF and CMML patient groups treated at 12 µg/kg, and separately for patients treated at 7 or 9 µg/kg. Efficacy results for MF patients will be presented separately for treatment-naïve and relapsed/refractory patients. For myelofibrosis patients, summaries will be displayed by DIPSS-Plus Risk for treatment naïve and for relapsed/refractory patients separately. For chronic myelomonocytic leukemia patients, summaries will be displayed by CMML Type for treatment naïve and for relapsed/refractory patients separately.

Treatment Naive (N=xx)	Relapsed/ Refractory (N=xx)	Overall (N=xx)

Efficacy summaries of investigator report response data for CMML patients will be presented separately for treatment-naïve and relapsed/refractory patients, with Stages 1 and 2 presented separately from Stage 3A patients.

Stages 1-2			Stage 3A		
Treatment Naive (N=xx)	Relapsed/ Refractory (N=xx)	Overall (N=xx)	Treatment Naive (N=xx)	Relapsed/ Refractory (N=xx)	Overall (N=xx)

Summaries of immunogenicity results will be presented by dose, separated by disease for patients treated at 12 µg/kg/day.

7 and 9 µg/kg/day	MF 12 µg/kg/day	CMML 12 µg/kg/day	Overall 12 µg/kg/day	Overall

3.6. Adjustments for Covariates

Due to small sample size, no formal statistical analyses that adjust for possible covariate effects are planned. Exploratory subgroup analyses may be performed for descriptive purposes.

3.7. Multiple Comparisons/Multiplicity

Multiplicity is not of concern for this study where Stages 1 and 2 were hypothesis generative cohorts. Stage 3A has a single primary efficacy endpoint: to analytically evaluate the clinical benefit derived from individual components, or variations thereof, of the MDS/MPN 2015 criteria. Additional endpoint analysis will be exploratory.

Type I error control will not be applied.

3.8. Subpopulations

Exploratory subgroup analyses may be performed for descriptive purposes.

3.9. Withdrawals, Dropouts, Loss to Follow-up

Patients who are withdrawn or discontinue from the study will not be replaced.

3.10. Missing, Unused, and Spurious Data

In general, there will be no substitutions made to accommodate missing data points. All data recorded on the eCRF will be included in data listings that will accompany the CSR.

For categorical variables, a category for missing data of the parameter will be presented. Percentages will be calculated from the total population, including patients with missing results.

When tabulating AE data, partial dates will be handled as follows. If the day of the month is missing, the onset day will be set to the first day of the month unless it is the same month and year as study treatment. In this case, in order to conservatively report the event as treatment-emergent, the onset date will be assumed to be the date of treatment. If the onset day and month are both missing, the day and month will be assumed to be January 1, unless the event occurred in the same year as the study treatment. In this case, the event onset will be coded to the day of

treatment in order to conservatively report the event as treatment-emergent. A missing onset date will be set up to the first day of treatment.

3.11. Visit Windows

It is expected that all visits should occur according to the protocol schedule. All data will be tabulated per the evaluation visit as recorded on the eCRF even if the assessment is outside of the visit window. In data listings, the relative day of all dates will be presented (see Section 3.2).

3.12. Timing of Analyses

[redacted] Efficacy data were not summarized at that time. The respective analyses were described in a separate Statistical Analysis Plan.

The final, end of study, analyses after database lock in support of the CSR will be supported from the analyses described in this document. Ongoing survival follow-up will be described in the CSR.

4. STUDY ANALYSES

4.1. Patient Disposition

Patient disposition will be tabulated as described in Section 3.5 and include the number screened, the number treated in total, the number in each patient population for analysis, the number who withdrew prior to completing treatment and reason(s) for withdrawal, and the number who withdrew prior to completing the study and reason(s) for withdrawal. The number of patients who received stem cell transplant since discontinuation will also be reported.

A data listing of study completion information including the reason for premature study withdrawal, if applicable, will be presented. A data listing of patient populations will also be presented.

4.2. Demographics and Baseline Characteristics

Demographics and baseline characteristics will be summarized as described in Section 3.5. Age, height, and weight will be summarized using descriptive statistics (number of patients, mean, SD, median, minimum, and maximum). The number and percentage of patients in each gender, race, ethnicity, and ECOG PS categories also will be presented.

Primary disease history will be summarized for MF and CMML. MF disease history will include DIPSS-plus score risk factors and profile, type of MF (primary, post polycythemia, post essential thrombocythemia), type of abnormal karyotype, mutation status, and time since diagnosis. CMML disease history will include diagnosis criteria, CMML type (CMML-1 or CMML-2), symptoms, cytogenic risk profile, monocular cell nodules in bone marrow biopsy (yes/no), mutations, prognostic scoring system molecular risk profile, and time since diagnosis. Baseline laboratory data, including monocytes, platelets, hemoglobin, white blood cell, lactate dehydrogenase, bone marrow blasts, peripheral blood blasts, will be summarized for all patients. Additionally, baseline MPN-SAF TSS, transfusion dependence status (yes/no) for red blood cells and platelets, spleen size by MRI/CT and by Physical Exam, the presence of splenomegaly at baseline (yes/no), and the presence of hepatomegaly at baseline (yes/no) will be summarized for all patients. SM disease history will be included in a data listing and will include diagnosis criteria, mutations, and molecular markers..

Receipt (yes/no) of prior systemic therapy, prior radiation therapy, and prior stem cell transplant (SCT) will be summarized. The number and type of prior systemic therapy for MPN will be presented. The reason why the patient was not considered an immediate candidate for SCT will also be summarized.

All analyses of demographic/baseline data will be performed using the Safety Population and select analyses will be repeated using the mITT Population.

Demographic, baseline, medical history, and disease history data for each patient will be provided in data listings.

4.3. Efficacy Evaluation

Efficacy analyses will be conducted using the mITT population for patients treated at 12 µg/kg/day. Supportive analysis of efficacy based on the Safety Population will be performed for patients treated at 12 µg/kg/day and patients treated at 7 or 9 µg/kg/day Efficacy summaries will be presented as described in Section 3.5. Data for all efficacy endpoints will be presented in by-patient listings.

4.3.1. Response Rate

Objective response rate and the corresponding 95% Clopper Exact CIs will be presented as the number and percentage of patients who achieve disease-specific complete response (CR) or partial response (PR) after treatment.

Best overall response will be presented as the number and percentage of patients who achieve disease-specific (see Section 8) CR, PR, stable disease (SD), progressive disease (PD) or who are not evaluable (NE) as their best response on study.

Swim lane plots will be produced for each disease presenting the time on treatment and in response.

The time from the date of first infusion of tagraxofusp to the date from when criteria are first met for CR or PR will be summarized.

Patients who remain in response per investigator and receive a stem cell transplant will be identified and summarized.

For MF patients, the number and percentage of patients who achieved clinical improvement, anemia, spleen, or symptom responses, molecular remission, or cytogenetic/molecular relapse will also be presented using IWG-MRT/ELN criteria (see Section 8.1 for disease specific response criteria).The time from date of first infusion of tagraxofusp to the date from when each response of clinical benefit criteria (as defined in Table 8-4 in Section 8.2) are first met will be summarized.

For CMML patients, the number and percentage of patients who achieved erythroid, platelet, neutrophil, spleen or symptoms responses will also be presented (see Section 8.2 for disease specific response criteria). The time from date of first infusion of tagraxofusp to the date from when each category of clinical benefit criteria are first met will be summarized.

4.3.2. Duration of Response

Duration of objective response is defined as the time from when the disease specific criteria (see Section 8) are first met for CR or PR until the date that the criteria for relapse (including SD, progressive disease [PD], or relapse) after CR/PR is met, or death. In the case that SD follows CR/PR and there is no evidence that response rebounds to CR/PR, duration of objective response will end at the time of first reduction of response to below CR/PR. For patients who receive SCT after CR/PR, DOR will include time to disease relapse post-transplant. Patients who are lost to follow-up or who do not relapse after objective response as of the analysis date will be censored on the latter of the date of last treatment with tagraxofusp or date of last disease

assessment recorded prior to the analysis date. If patient starts other anticancer therapies (with the exception of SCT), they will be censored at the latter of the date of last treatment with tagraxofusp or date of last disease assessment recorded that occurred prior to the start of the new therapy.

The distribution of DOR will be estimated by Kaplan-Meier methodology and the 25th percentile, median, and 75th percentile of time to event, number and percentage of events and censored observations, and appropriate 95% CIs will be presented.

For MF patients, the median duration of response for each category of clinical benefit, including clinical improvement, anemia, spleen, or symptom responses, molecular remission, or cytogenetic/molecular relapse, will be derived similar to the duration of objective response. For CMML patients, the median duration of response for each category of clinical benefit, including erythroid, platelet, neutrophil, spleen or symptoms responses, will be derived similar to the duration of objective response.

4.3.3. Progression-Free Survival

Progression-free survival is defined as the time from the date of first infusion of tagraxofusp to the date of PD or death from any cause, whichever occurred first. For patients who receive SCT, PFS will include time to PD or death post-transplant. Patients who do not progress and are still alive at the time of analysis will be censored on the date of last disease assessment recorded prior to the analysis date. If patient starts other anticancer therapies (with the exception of SCT), they will be censored at the latter of the date of last treatment with tagraxofusp or date of last disease assessment recorded that occurred prior to the start of the new therapy. The distribution for PFS will be estimated by Kaplan-Meier methodology and the 25th percentile, median, 75th percentile, number and percentage of events and censored observations, and appropriate 95% CIs will be presented.

4.3.4. Overall Survival

Overall survival is defined as the time from the date of first infusion of tagraxofusp to the date of death from any cause. Patients still alive or lost to follow-up at the time of the analysis will be censored on the last date known to be alive prior to the analysis date, as determined by in-person visit or telephone contact.

Duration of follow-up is defined as the time from the date of first infusion of tagraxofusp to the last date known alive. Patients who died will be censored on the date of death.

The overall distributions for OS and follow-up durations will be estimated by Kaplan-Meier methodology in a similar manner to PFS.

4.3.5. Spleen Volume

Spleen size by MRI/CT will be summarized over time for all patients with baseline splenomegaly. The number and percentage of patients with ≥ 10%, ≥ 25 %, and ≥ 35% reduction in spleen volume from baseline will be summarized over time.

The maximum percent reduction from baseline in spleen size across all post-baseline visits will be summarized. A waterfall plot of the maximum percent reduction in spleen volume, spleen volume at Cycle 4 Day 21 visit, and spleen volume at Cycle 7 Day 28 will be created.

4.3.6. Tumor Symptom Score

Symptom scores will be summarized over time for all patients. The number and percentage of patients with ≥ 25% and ≥ 50% reduction in MPN-SAF TSS from baseline will be summarized over time.

The maximum percent reduction from baseline in score across all post-baseline visits will be summarized. A waterfall plot of the maximum percent reduction in score, score at Cycle 4 Day 21 visit, and score at Cycle 7 Day 28 will be created.

4.3.7. Other Assessments of Disease Response

The proportion of subjects with ≥ 35% reduction in spleen size from baseline and ≥ 50% reduction in score from baseline will be summarized over time. The proportion of subjects with an anemia response for MF patients and with erythroid response for CMML patients will also be summarized over time.

For patients who were transfusion-dependent at baseline, the number and percentage of patients who achieved transfusion independence following treatment start will be summarized for any 12-week period (at any time, at C4D21-12 weeks and C7D28-24 weeks). Transfusion dependency will be for pRBCs, platelets, and both, as defined in Section 1.2.5.1. The rate of transfusions will also be summarized at any time post-baseline, from pre-baseline to baseline, from baseline to Week 12, and from Week 12 to Week 24.

For patients who were Red Blood Cell Transfusion Independent at Baseline, the number and percentage of patients becoming transfusion dependent following treatment start will be summarized for any 12-week period (at any time, at C4D21-12 weeks and C7D28-24 weeks). Transfusion dependency will be summarized pRBC as defined in Section 1.2.5.1.

4.3.8. Laboratory Efficacy Results

4.3.8.1. Bone Marrow Blasts

Bone marrow blasts will be summarized over time for all patients. In case of different methods used to assess bone marrow at the same visit (i.e aspirate and biopsy) but only one method was used at baseline, the method used at baseline will be considered for comparison. In case both methods are used in all visits, the larger blast results between biopsy results and aspirate results for a subject will be presented. In case only one method is used in a specific visit, it will be considered regardless the method used at baseline.

The number and percentage of patients with ≥ 50% reduction in bone marrow blasts from baseline and the number and percentage of patients with < 5 % blasts will be summarized at any time for CMML patients with baseline blasts >=5%.

The maximum percent reduction from baseline and the maximum increase from baseline across all post-baseline visits will be summarized for CMML patients. A waterfall plot of the maximum percent reduction in blasts, blasts at Cycle 4 Day 21 visit, and blasts at Cycle 7 Day 28 will be created.

Mean and standard error will be plotted over time for CMML.

4.3.8.2. Peripheral Blood Blasts

Peripheral blood blasts will be summarized over time for all patients.

The number and percentage of patients with ≥ 50% reduction in peripheral blood blasts from baseline and the number and percentage of patients with 0 % blasts will be summarized at any time for CMML patients with baseline blasts >=0%.

The maximum percent reduction from baseline and the maximum increase from baseline across all post-baseline visits will be summarized for CMML patients. A waterfall plot of the maximum percent reduction in blasts, blasts at Cycle 4 Day 21 visit, and blasts at Cycle 7 Day 28 will be created.

Mean and standard error will be plotted over time for CMML.

4.3.8.3. Absolute Monocyte Count

Absolute monocyte counts will be summarized over time for all CMML patients.

The number and percentage of patients with ≥ 50% reduction in absolute monocyte count from baseline and the number and percentage of patients with counts within normal limits will be summarized at any time.

The maximum percent reduction from baseline and the maximum increase from baseline across all post-baseline visits will be summarized. A waterfall plot of the maximum percent reduction in counts, counts at Cycle 4 Day 21 visit, and counts at Cycle 7 Day 28 will be created.

Mean and standard error will be plotted over time for CMML.

4.3.8.4. White Blood Cell Count

White blood cell counts will be summarized over time for all patients.

The number and percentage of patients with ≥ 50% reduction in white blood cell count from baseline and the number and percentage of patients with counts within normal limits will be summarized at any time for CMML patients.

The maximum percent reduction from baseline and the maximum increase from baseline across all post-baseline visits will be summarized for CMML patients. A waterfall plot of the maximum percent reduction in counts, counts at Cycle 4 Day 21 visit, and counts at Cycle 7 Day 28 will be created.

Mean and standard error will be plotted over time.

4.3.8.5. Hemoglobin

Hemoglobin will be summarized over time for all patients.

The number and percentage of patients with ≥ 20 g/L increase in hemoglobin from baseline and the number and percentage of patients with results within normal limits will be summarized at any time for CMML patients.

The maximum percent reduction from baseline and the maximum increase from baseline across all post-baseline visits will be summarized for CMML patients. A waterfall plot of the maximum increase in hemoglobin, hemoglobin at Cycle 4 Day 21 visit, and hemoglobin at Cycle 7 Day 28 will be created.

Mean and standard error will be plotted over time.

4.3.8.6. Platelet Count

Platelet counts will be summarized over time for all patients.

The number and percentage of patients with ≥ 30 x 10^9/L increase in platelet count from baseline, the number and percentage of patients with ≥ 100% increase in platelet count, and the number and percentage of patients with results within normal limits will be summarized at any time for CMML patients.

The maximum percent reduction from baseline and the maximum increase from baseline across all post-baseline visits will be summarized for CMML patients. A waterfall plot of the maximum increase in counts, counts at Cycle 4 Day 21 visit, and counts at Cycle 7 Day 28 will be created.

Mean and standard error will be plotted over time.

4.4. Safety Analyses

Safety analyses will be conducted using the Safety population and presented as described in Section 3.5.

4.4.1. Study Drug Exposure

Duration of study drug exposure will be calculated as the number of days and number of cycles patients were administered study drug.

$$Duration\ of\ Study\ Drug\ Exposure = (Date\ of\ last\ dose - Date\ of\ first\ dose) + 1$$

Number of cycles of study drug initiated while the patient is on study will be summarized overall. Dose modifications, including the number of patients with dose interruptions or dose

reductions, will be summarized by cycle. Total dose administered will be summarized overall and number of doses administered will be summarized by cycle.

Relative Dose Intensity will be computed using the following definition and summarized by disease by trial stage:

$$Relative\ Dose\ Intensity\ (\%) = \frac{Sum(Cumulative\ Actual\ Dose\ Received\ by\ Cycle)}{Sum(Planned\ Dose\ to\ be\ Administered\ by\ Cycle)} \times 100$$

Exposure to tagraxofusp will be presented in a data listing based on Safety Population.

4.4.2. Adverse Events

All AEs will be coded using the MedDRA coding system version 19.0 and displayed in tables and data listings using system organ class (SOC) and preferred term (PT) based on Safety Population.

Analyses of AEs will be performed for those events that are considered treatment-emergent, where treatment-emergent is defined per protocol as any AE with onset after the first administration of tagraxofusp through 30 days after the last dose of tagraxofusp, or any event that was present at baseline but worsened in intensity or was subsequently considered drug-related by the Investigator through 30 days after the last dose of tagraxofusp.

Tabulations of the number and percentage of patients with any treatment-emergent AE (TEAE), common AEs (defined as occurring in ≥10% of the study sample [overall for summaries by dose, within each disease for summaries by disease]), with any TEAE assessed by the Investigator as related to treatment, with any TEAE with severity ≥ Common Terminology Criteria for Adverse Events (CTCAE) Grade 3, with any serious adverse event (SAE), with any AE leading to discontinuation of study treatment, with any AE to dose reduction, with any AE leading to dose interruption, with any AE leading to death, with any AE leading to death within 30 days, and with any AE leading to death within 60 days will be summarized as described in Section 3.5. An overall summary of the incidence of TEAEs will also be provided.

In these tabulations, each patient will contribute only once (ie, the most related occurrence or the most intense occurrence) to each of the incidence rates in the descriptive analysis, regardless of the number of episodes. Treatment-emergent AEs (TEAEs) are summarized by patient incidence rates; therefore, in such tabulations, a patient contributes only once to the count for a given SOC or PT. For tabulations that include classification by relationship to study treatment, AEs with missing relationship will be considered related to study drug.

By-patient tabulations will also be presented for patients who died, SAEs, AEs leading to study drug discontinuation, AEs with CTCAE Grade 3, and AEs with CTCAE Grade 4.

Analysis of AEs over time will be conducted by presentation of TEAE incidence by cycle. For this analysis, incidence per cycle will be based on the onset of a particular PT event within the cycle, where onset is indicated by first or repeat occurrence or increase in severity of the event within that cycle. This analysis will be conducted by maximum CTCAE Grade for each PT. AEs over time will also be presented by cycle for the incidence of TEAE leading to treatment discontinuation, dose reduction, and dose interruption by SOC and PT.

AEs of special interest (AESIs) will be determined using MedDRA version 19.0 standardized medical queries (SMQs), high-level terms, or PTs, as follows:

- Possible hypersensitivity events, based on the SMQ Hypersensitivity (broad search).
- Vascular capillary leak syndrome (CLS), in addition to PTs of hypoalbuminaemia, blood albumin decreased, and proteinuria.
- Visual acuity, based on the MedDRA PTs related to vision changes.
- Post-transplantation veno-occlusive disease, based on the MedDRA PTs pulmonary veno-occlusive disease and veno-occlusive liver disease.
- Possible drug-induced liver injury events, based on the SMQ drug-related hepatic disorders – comprehensive search (broad search).

Summary tables of AESIs will report AESI by PT, causality, severity grade, seriousness, and whether AESI led to discontinuation of study treatment. AESIs will also be included in data listings.

4.4.2.1. Capillary Leak Syndrome

In addition to the summaries described in Section 4.4.2, summaries will be provided that include:

- Summary of number of CLS events, number of CLS events by CTCAE Grade, and time to first onset and time to resolution of CLS events;
- Summary incidence of albumin infusions and timing of infusions in relation to CLS onset;
- Summary incidence of steroids used
- Summary incidence of recurring CLS
- Summaries of CLS leading to treatment discontinuation, dose reduction, and dose interruption

4.4.3. Laboratory Data

Clinical laboratory values will be expressed in Système International (SI) units and tabulations will be based on Safety Population.

The actual value and change from baseline (Day 1) to each on-study evaluation through cycle 6 will be summarized for each clinical laboratory parameter, including hematology, clinical chemistry, coagulation, and urinalysis. In the event of repeat values, the last non-missing value closest to the planned study day/time will be used. Box plots will also be presented.

Shift tables of change in CTCAE grade of laboratory parameters from baseline to worst value and from baseline to last value on study will be presented for each disease. Both scheduled and unscheduled visits will be included in shift tables.

To assess for possible drug-induced liver injury (DILI) (FDA 2009), alanine aminotransferase, aspartate aminotransferase, and total bilirubin results will be listed for patients that meet Hy's law criteria. Hy's Law case candidates are defined as subjects with peak ALT or AST >3 × upper limit of normal (ULN) and total bilirubin >2 × ULN within ± 7 days.

All laboratory data will be provided in by-patient data listings.

By-patient listings will also be presented for all laboratory values with CTCAE Grade ≥3 and for all clinically significant laboratory values with CTCAE Grade ≥3.

4.4.4. Vital Signs and Physical Examination

The actual value and change from baseline to each on-study evaluation and to the last on study evaluation will be summarized for vital signs based on Safety Population.

A summary table of the number and percent of patients with treatment-emergent potentially clinically significant (PCS) vital signs parameters will be tabulated based on the following criteria:

	PCS – Low If:			**PCS – High if:**		
Variable Name	**Observed Value is:**	**AND**	**Decrease from Baseline is:**	**Observed Value is:**	**AND**	**Decrease from Baseline is:**
Systolic Blood Pressure	<90 mmHg		≥20 mmHg	>180 mmHg		≥20 mmHg
Diastolic Blood Pressure	<50 mmHg		≥10 mmHg	>105 mmHg		≥10 mmHg
Heart Rate	<50 bpm		≥25 bpm	>120 bpm		≥15 mmHg

All tables summarizing vital sign measurements only include visits in which at least 10% of the analysis population had measurements. Vital sign measurements will be presented for each patient in a data listing.

All physical examination findings will be presented in data listings.

4.4.5. Electrocardiogram

Actual value and change from baseline will be summarized by visit over time for heart rate, QT interval, and QTc interval; continuous summary statistics will be based on the mean ECG values at each timepoint, if appropriate. All tables summarizing ECG measurements only include visits in which at least 10% of the analysis population had measurements.

Based on the 12-lead ECG results, the number and percentage of patients whose mean QTc value at any point meetings any of the following categories will be summarized:'

- >450 msec
- >480 msec
- >500 msec
- increase from baseline >30 msec
- increase from baseline >60 msec

Electrocardiogram and echocardiogram/MUGA data for each patient will be provided in data listings.

4.4.6. Concomitant Medications

Concomitant medications will be coded using the WHO Drug Dictionary, version September 2016. Results will be tabulated by anatomic therapeutic class (ATC) and PT.

Concomitant medications will be tabulated where any medications that were not discontinued prior to first dose of study drug and those started after the first dose of study treatment will be included. If an end date is missing or the medication is ongoing, the medication will be considered concomitant.

The use of concomitant medications and subsequent treatment(s) will be included in by-patient data listings.

Prior systemic therapies for MPN will also be summarized by ATC and PT.

4.5. Summary of Immunogenicity Parameters

Immunogenicity parameters will be summarized as described in Section 3.5.

Analyses will be conducted in accordance with the principles outlined in (Shankar, et al. 2014) and will include:

- Incidence of patients with positive anti-drug antibodies, specific anti-hIL-3 antibodies, and neutralizing antibodies will be assessed by cycle uniquely through cycle 6 and cycles >6, collectively. Incidence will be summarized by number and percent of patients with positive antibody results.
- Mean fold increase from baseline and geometric mean titer are tabulated by scheduled assessments. Summary statistics will include mean, standard deviation, median, min, and max.
- Incidence of patients with a $\geq$ 10 fold increase in anti-drug antibody titer and/or a $\geq$ 4 fold increase in anti-hIL-3 antibody titer will be tabulated.

Immunogenicity parameters will be included in by-patient data listings.

4.6. Summary of Pharmacokinetics

Individual plasma concentrations-time profiles of free Tagraxofusp will be listed by study stage, subject, study cycle/day, dose, ADA titer and disease. All concentration results reported as BLQ will be labeled as such in the data listings.

Individual plasma concentration-time profiles of free Tagraxofusp will be presented graphically (linear and semi-log scale) and overlaid: (i) by study cycle/day and dose, (ii) by study cycle/day, dose and baseline ADA titer, (iii) by study cycle/day, dose and study stage, and (iv) by study

cycle/day, dose and disease. BLQ concentration results will be treated as missing in the semi-log plots.

As data permit, the following descriptive summary statistics of plasma concentrations vs. time will be reported: Geometric N, Geometric mean, geometric SD, geometric %CV, N, arithmetic mean, SD, %CV, median, min and max.

Descriptive summary statistics of plasma concentrations vs. time will be presented in tabular form: (i) by study cycle/day and dose, (ii) by study cycle/day, dose and baseline ADA titer, (iii) by study cycle/day, dose and study stage, and (iv) by study cycle/day, dose and disease.

Mean plasma concentration-time profiles of free Tagraxofusp will be presented graphically (linear and semi-log scale) and overlaid: (i) by study cycle/day and dose, (ii) by study cycle/day, dose and baseline ADA titer, (iii) by study cycle/day, dose and study stage, and (iv) by study cycle/day, dose and disease.

Individual PK parameters of free Tagraxofusp will be listed by study stage, subject, study cycle/day, dose, ADA titer and disease. Details concerning derivation of the parameter values will be included.

As data permit, the following descriptive summary statistics of PK parameters will be reported: Geometric N, Geometric mean, geometric SD, geometric %CV, N, arithmetic mean, SD, %CV, median, min and max.

Descriptive summary statistics of PK parameters will be presented in tabular form: (i) by study cycle/day and dose, (ii) by study cycle/day, dose and baseline ADA titer, (iii) by study cycle/day, dose and study stage, and (iv) by study cycle/day, dose and disease.

Distribution of C_{max} and AUC_{last} parameters will be presented graphically (boxplots) and overlaid: (i) by study cycle/day and dose, (ii) by study cycle/day, dose and study stage, (iii) by study cycle/day, dose and disease, and (iv) by study cycle/day, dose and baseline ADA titer.

5. CHANGES TO PLANNED ANALYSES

Planned analyses for the activity of tagraxofusp in specific sub-populations of patients with MF that are of interest will not be performed.

An interim analysis for Stages 1-2 and Stage 3A data was planned to be performed after enrollment in Stage 3A is complete. Analyses were to include all analyses described in this SAP, as well as correlative and Bayesian analyses. These analyses would serve as the primary evidence of efficacy for regulatory decision-making in the planned confirmatory cohort (Stage 3B).

After an ongoing review of safety and efficacy data, Stemline Therapeutics decided to not continue enrollment into Stage 3A and not start enrollment into Stage 3B. Therefore, Bayesian and correlative analyses will not be performed. Formal descriptive analyses will be performed after all patients have exited the study and the database is locked.

Time to responses, including overall response and clinical benefit subcategories, duration of response of clinical benefit subcategories, and bridge to stem cell transplant summaries were added. The analyses of laboratory results related to blast percentages, white blood cells, hemoglobin, monocytes, and platelets were also added.

Further analyses of transfusion dependency and transfusion rates over time were not pre-specified but added to the plan.

6. REFERENCES

Brunetti L, Di Battista V, Venanzi A, et al. Blastic plasmacytoid dendritic cell neoplasm and chronic myelomonocytic leukemia: a shared clonal origin. Leukemia. 2017;31(5):1238-1240.

Chauhan D, Ray A, Das DS, et al. Effect of a novel agent, SL-401, targeting interleukin-3 receptor (IL-3R) on plasmacytoid dendritic cell (pDC)-induced myeloma cell growth and osteolytic bone disease. Proc Am Soc Clin Oncol 2014; 32(15S):563s (A8599).

Cheson BD, Greenberg PL, Bennett JM, et al. Clinical application and proposal for modification of the international working group (IWG) response criteria in myelodysplasia. Blood 2006; 108(2): 419-425.

Facchetti F, Cigognetti M, Fisogni S. Neoplasms derived from plasmacytoid dendritic cells. Mod Pathol. 2016;29(2):98-111.

FDA. Guidance for Industry: Drug-Induced Liver Injury: Premarketing Clinical Evaluation. 2009.

Ji P, Peterson LC. Plasmacytoid dendritic cells in chronic myelomonocytic leukemia. Blood. 2014;123(21):3220.

Krishnan A, Li B, Pagane M, et al. Evaluation of Combination Tagraxofusp (SL-401) and Hypomethylating Agent (HMA) Therapy for the Treatment of Chronic Myelomonocytic Leukemia (CMML). Blood. 2018 132:1809.

Lucas N, Duchmann M, Rameau P, et al. Biology and prognostic impact of clonal plasmacytoid dendritic cells in chronic myelomonocytic leukemia. Leukemia. 2019;33(10):2466-2480.

Naresh KN, Pavlu J. Plasmacytoid dendritic cell nodules in bone marrow biopsies of chronic myelomonocytic leukemia. Am J Hematol. 2010;85(11):893.

Pophali P, Horna P, Lasho TL, et al. Splenectomy in patients with chronic myelomonocytic leukemia: Indications, histopathological findings and clinical outcomes in a single institutional series of thirty-nine patients. Am J Hematol. 2018;93(11):1347-1357.

Ricci C, Fermo E, Corti S, et al. RAS mutations contribute to evolution of chronic myelomonocytic leukemia to the proliferative variant. Clin Cancer Res. 2010;16(8):2246-56.

Savona MR, Malcovati L, Komrokji R, et al. An international consortium proposal of uniform response criteria for myelodysplastic/myeloproliferative neoplasms (MDS/MPN) in adults. Blood. 2015;125(12):1857-65.

Shankar G, Arkin S, Cocea L, Devanarayan V, Kirshner S, Kromminga A, et al. Assessment and reporting of the clinical immunogenicity of therapeutic proteins and peptides-harmonized terminology and tactical recommendations. *AAPS J.* 2014;16(4):658-73.

Shen Q, Ouyang J, Tang G, et al. Flow cytometry immunophenotypic findings in chronic myelomonocytic leukemia and its utility in monitoring treatment response. Eur J Haematol. 2015;95(2):168-76.

Tefferi A, Cervantes F, Mesa R, et al. Revised response criteria for myelofibrosis: international working group – myeloproliferative neoplasms research and treatment (IWG-MRT) and European LeukemiaNet (ELN) consensus report. Blood 2013B; 122:1395-98.

Vermi W, Facchetti F, Rosati S, et al. Nodal and extranodal tumor-forming accumulation of plasmacytoid monocytes/interferon-producing cells associated with myeloid disorders. Am J Surg Pathol. 2004 May;28(5):585-95.

Vermi W, Soncini M, Melocchi L, et al. Plasmacytoid dendritic cells and cancer. J Leukoc Biol. 2011;90:681-690.

Watkins PB, Seligman PJ, Pears JS, Avigan MI, and Senior JR. Using controlled clinical trials to learn more about acute drug-induced liver injury. *Hepatology.* 2008;48(5):1680-9.

7. STUDY FLOW CHARTS

The following Study Flow Charts are from Protocol Amendment 9.

Table 7-1 Stage 2: Schedule of Events for Screening and Cycles 1-4

Tests and Observations	Day -28 to 0	Days 1-3 (Up to Day 10 if Infusion(s) Held) Tagraxofusp Treatment		Cycle 1 Day 8±3[1]	Day 21±3[2]	Day 28±3, Then Every 7±3 days *(Only If Delayed End of Cycle: for Toxicity Resolution)*
	Screening	*Pre-Infusion*	*Tagraxofusp Infusion*		*End of Cycle*[2]	
Informed consent form[3]	X					
Inclusion/exclusion criteria[4]	X					
Medical history	X					
ECOG performance status	X	X (Infusion 1, C1)			X	X
Physical examination including assessment of hepatomegaly and splenomegaly	X			X	X	
Pregnancy test[5]	X					
Vital signs and weight[6]	X	X	X	X	X	X
12-lead ECG[7]	X	X (MF Only: Infusion 1) (C1, C2)	X (MF Only: Infusion 1) C1, C2)			
MUGA scan or 2-D Echocardiogram[8]	X					
Hematology[9]	X	X		X	X	X
Serum chemistry[10]	X	X		X	X	X
Coagulation parameters: PT/INR, aPTT[11]	X	X		X	X	X
Urinalysis[12]	X			X	X	X
Peripheral blood for flow cytometry	X	X (Infusion 1)		X		
Bone marrow aspiration/biopsy[13]	X				X C1 and C4[14]	
MRI or CT scan of abdomen[15]	X				X C1 and C4[14]	

Tests and Observations	Day -28 to 0	Days 1-3 (Up to Day 10 if Infusion(s) Held) Tagraxofusp Treatment		Cycle 1 Day 8±3[1]	Day 21±3[2]	Day 28±3, Then Every 7±3 days *(Only If Delayed End of Cycle: for Toxicity Resolution)*
	Screening	*Pre-Infusion*	*Tagraxofusp Infusion*		*End of Cycle*[2]	
Pharmacokinetic sampling (patient with MF only)[16]		X (Infusion 1) (C1, C2)				
Immunogenicity sampling[17]		X (Infusion 1)			X	
Cytogenetic and molecular genetic testing	X				X C1 and C4	
Administration of premedications[18]		X				
Tagraxofusp administration[19]			X			
MPN-SAF TSS evaluation[20]	X				X	
Tumor response assessment[21]	X				X C1 and C4	
Vision assessment	X				X	
Prior/concomitant medications and therapies	X	X	X	X	X	X
AE and SAE monitoring	X	X	X	X	X	X

AE = adverse event; aPTT = activated partial thromboplastin time; CT = computed tomography; ECG = electrocardiogram; ECOG = Eastern Cooperative Oncology Group; INR = international normalized ratio; IV = intravenous; MPN-SAF TSS = Myeloproliferative Neoplasm Symptom Assessment Form Total Symptom Score; MUGA = multigated acquisition; PK = pharmacokinetics; PT = prothrombin time; SAE = serious adverse event.

1 Day 8 visit is required for C1 only and must be performed at the site.

2 The end-of-cycle evaluations (Day 21 or thereafter) may also serve as the pre-infusion evaluations for the subsequent cycles; with the exception of vital signs, these assessments do not need to be duplicated on successive days unless there is an abnormality or other clinically relevant reason for repeat evaluation.

3 Refer to protocol Section 16.3 for details.

4 Refer to protocol Section 7 (study populations) for details.

5 Urine or serum pregnancy test must be performed within 1 week prior to treatment for WOCBP.

6 Height will be measured at screening only; weight does not need to be measured more than once per day, and should be measured pre-infusion on treatment days. Vital signs should be performed after patient is sitting for 3 to 5 minutes. During dosing period, vital signs should be taken immediately prior to infusion, at 0 (i.e., immediately after completion of infusion), and at 30 and 60 minutes post-infusion.

7 All patients will have a 12-lead ECG performed at the screening visit. For MF patients only, during the days when patients are undergoing PK sampling (Cycles 1 and 2, Infusion 1), an ECG will be performed at 3 distinct time points (triplicates) within 5 minutes (±5 minutes) prior to each PK sample collection pre-infusion and at 30 and 60 minutes post-infusion.

8 A MUGA scan or 2-D ECHO to quantify LVEF must be completed within 28 days prior to start of first cycle of study drug.

9 To be collected prior to tagraxofusp infusion if during dosing period. Hematology includes WBC count with differential, RBC count, hematocrit, Hb, platelet count and immature myeloid cells (including blasts + myelocytes +metamyelocytes + promyelocytes + nucleated red blood cells).

10 To be collected prior to tagraxofusp infusion if during dosing period. Serum chemistry incudes electrolytes and additional parameters (equivalent to Chem-20): ALT, albumin, ALP, AST, bicarbonate, bilirubin, BUN, calcium, magnesium, chloride, creatinine, glucose, LDH, phosphate, potassium, sodium, total protein, uric acid, and CPK. See protocol Appendix Section 18.3 for administration of albumin if serum albumin decreases to < 3.0 g/dL (< 30 g/L) during treatment days or in the immediate post-treatment period.

11 In lieu of PT, INR may be measured.

12 Urinalysis includes appearance, color, pH, specific gravity, ketones, leukocytes, protein, glucose, bilirubin, urobilinogen, and occult blood. Dipstick is acceptable.

13 Morphology and WBC differential/blast count on aspirate. Baseline must be performed within 28 days prior to the first administration of tagraxofusp. Subsequent bone marrow aspirates and biopsies will be performed at the end of C1, C4, and then every 12 weeks (±7 days). Bone marrow evaluation should also be performed at End of Treatment including patients who discontinue study therapy prior to completion of C4; if a bone marrow evaluation was performed within 4 weeks prior to the end of treatment, a repeat evaluation does not need to be performed at this time, unless clinically indicated.

14 Bone marrow and imaging studies are to be repeated at the end of C1 and C4 and then every 12 weeks (±7 days) thereafter.

15 A baseline MRI or CT scan must be performed within 28 days prior to the first administration of tagraxofusp. Subsequent imaging studies will be performed at the end of C1 and C4 and then every 12 weeks (±7 days) until there is evidence of relapsed or progressive disease. Abdominal scan is required; evaluation of chest/pelvis may be obtained at the Investigator's discretion. Scan must include measurement of spleen and liver volume. Imaging studies should also be performed at End of Treatment, including patients who discontinue study therapy prior to completion of C4; if a imaging was performed within 4 weeks prior to the end of treatment, a repeat evaluation does not need to be performed at this time, unless clinically indicated. Whenever clinically feasible, it is strongly requested that the same imaging modality (MRI or CT) at both baseline and subsequent timepoints be utilized for any given patient throughout their evaluation on-study.

16 Peripheral blood (plasma) samples (6 mL each) will be collected from patients with MF only immediately prior to the start of the infusion of tagraxofusp, then at 0 (i.e., upon completion of infusion), 15, 30, 45, 60, 90 120, 180, and 240 minutes after completion of the infusion during Infusions 1 (i.e., Day 1) in C1 and C2.

17 Peripheral blood (serum) samples (10 mL) will be collected for the detection of tagraxofusp reactive antibodies on Day 1 (pre-infusion) and Day 21 (the end of cycle collection may also serve as the pre-infusion collection for the subsequent cycle).

18 Premedications administered 60 minutes (±15 minutes) before tagraxofusp treatment: Diphenhydramine 50 mg IV (or an equivalent dose of another H_1-histamine antagonist); acetaminophen 650 mg orally (or equivalent dose of paracetamol); methylprednisolone 50 mg IV (or an equivalent dose of another corticosteroid); famotidine 20 mg IV (or an equivalent dose of another H_2-histamine antagonist).

19 Following treatment with premedication, tagraxofusp will be administered as a 15-minute infusion once daily for the first 3 consecutive days of a 21-day cycle. Individual tagraxofusp infusions may be delayed to allow for toxicity resolution, but all 3 infusions should be completed within 10 days.

20 Patients with MF should complete the MPN-SAF TSS (10-question evaluation concerning symptoms/well-being during the prior week) during screening and in every treatment cycle, prior to receiving therapy on the subsequent cycle. The MPN-SAF TSS should also be completed at End of Treatment.

21 Tumor assessments are to be performed during screening, at the end of C1 and C4, then every 12 weeks (±7 days) and at End of Treatment. These include the IWG-MRT/ELN 2013 consensus report for MF and the IWG MDS 2016 consensus report for CMML. These are detailed in protocol Section 11.12 and Section 18.1

Table 7-2 Stage 2: Study Events Schedule for Cycle 5 and Beyond in Stage 2

	Cycle 5 and Beyond				Cycles 5 and Beyond[2] 12 Weeks (i.e., every 3 cycles) ±7 Days	End-of-treatment	Safety: Through 30 Days After Last Infusion	Survival: Every 90 Days After Last Infusion
	Days 1-3 (Up to Day 10 if Infusion(s) Held) Tagraxofusp Treatment		Day 28±3[1]					
Procedures	***Pre-Infusion***	***Infusion***	***End of Cycle***[1]					
Concomitant medications/therapies	X		X			X		
AE and SAE monitoring	X	X	X		X	X	X	
ECOG performance status			X			X		
Physical examination including assessment of hepatomegaly and splenomegaly			X			X		
Vital signs and weight[3]	X	X	X			X		
Hematology[4]	X		X			X		
Serum chemistry[5]	X		X			X		
Coagulation parameters: PT/INR, aPTT[6]	X		X			X		
Urinalysis[7]			X					
Peripheral blood for flow cytometry	X (Infusion 1)					X		
Bone marrow aspiration/biopsy[8]					X[9]	X		
MRI or CT scan of abdomen[10]					X[9]	X		
Immunogenicity sampling[11]			X			X	X	X[12]
Cytogenetic and molecular genetic testing					X	X		
Administration of premedications[13]	X							
Administration of Tagraxofusp[14]		X						
MPN-SAF TSS evaluation[15]			X			X		
Vision assessment			X			X		
Tumor response assessment[16]					X[10]	X		

Procedures	Cycle 5 and Beyond: Days 1-3 (Up to Day 10 if Infusion(s) Held) Tagraxofusp Treatment: *Pre-Infusion*	Cycle 5 and Beyond: Days 1-3 (Up to Day 10 if Infusion(s) Held) Tagraxofusp Treatment: *Infusion*	Cycle 5 and Beyond: Day 28±3[1]: *End of Cycle*[1]	Cycle 5 and Beyond	Cycles 5 and Beyond[2] 12 Weeks (i.e., every 3 cycles) ±7 Days	End-of-treatment	Safety: Through 30 Days After Last Infusion	Survival: Every 90 Days After Last Infusion
Long-term follow-up[12]							X	X

AE = adverse event; aPTT = activated partial thromboplastin time; CT = computed tomography; ECG = electrocardiogram; ECOG = Eastern Cooperative Oncology Group; INR =international normalized ratio; MPN-SAF TSS = Myeloproliferative Neoplasm Symptom Assessment Form Total Symptom Score; PT = prothrombin time; SAE = serious adverseevent.

1 The end-of-cycle evaluations (Day 28 or thereafter) may also serve as the pre-infusion evaluations for the subsequent cycles; with the exception of vital signs, these do notneed to be duplicated on successive days unless there is an abnormality or other clinically relevant reason for repeat evaluation.

2 Disease evaluation via bone marrow aspirate/biopsy and/or imaging studies (when relevant) will occur every 12 weeks (±7 days). The frequency of bone marrow evaluationmay be reduced (obtained less frequently) if peripheral blood, clinical and/or imaging assessments indicate clinical stability, as per the Investigator's discretion in consultation with the Medical Monitor. The frequency of imaging assessments may also be reduced if clinical/laboratory findings indicate stability, as per the Investigator'sdiscretion in consultation with the Medical Monitor.

3 Height will be measured at screening only; weight does not need to be measured more than once per day, and should be measured pre-infusion on treatment days. Vital signsshould be performed after patient is sitting for 3 to 5 minutes. If during dosing period, vital signs should be taken immediately prior to infusion, at 0 (i.e., immediately after completion of infusion), and at 30 and 60 minutes post-infusion.

4 To be collected prior to tagraxofusp infusion if during dosing period. Hematology includes WBC count with differential, RBC count, hematocrit, hemoglobin, platelet countand immature myeloid cells (including blasts + myelocytes +metamyelocytes + promyelocytes + nucleated red blood cells).

5 To be collected prior to tagraxofusp infusion if during dosing period. Serum chemistry incudes electrolytes and additional parameters (equivalent to Chem-20): ALT, albumin,alkaline phosphatase, AST, bicarbonate, bilirubin, BUN, calcium, magnesium, chloride, creatinine, glucose, LDH, phosphate, potassium, sodium, total protein, uric acid, and CPK. See protocol Section 18.3 for administration of albumin if serum albumin decreases to < 3.0 g/dL (< 30 g/L) during treatment days or in the immediate post-treatment period.

6 In lieu of PT, INR may be measured.

7 Urinalysis includes appearance, color, pH, specific gravity, ketones, leukocytes, protein, glucose, bilirubin, urobilinogen, and occult blood. Dipstick is acceptable.

8 Morphology and WBC differential /blast count on aspirate. In general, bone marrow aspirates and biopsies and peripheral blood samples will be performed at the end of C1and C4 and then every 12 weeks (±7 days) thereafter. Bone marrow evaluation should also be performed at End of Treatment; if a bone marrow evaluation was performed within 4 weeks prior to the end of treatment, a repeat evaluation does not need to be performed at this time, unless clinically indicated.

9 Bone marrow and imaging studies are to be repeated at the end of C1 and C4 and then every 12 weeks (±7 days) thereafter.

10 Imaging studies will be performed at the end of C1 and C4 and then every 12 weeks (±7 days) thereafter until there is evidence of relapsed or progressive disease. Abdominal scan is required; evaluation of chest/pelvis may be obtained at the Investigator's discretion. Scan must include measurement of spleen and liver volume. Subsequent to Week 24, for patients receiving ongoing therapy (or those who discontinue therapy but are without evidence of PD and continue to be evaluated in ongoing follow-up) the frequencyof imaging studies may be reduced (obtained less frequently) if peripheral blood and/or clinical assessments indicate clinical stability,

as per the Investigator's discretion, in consultation with the Medical Monitor. Imaging studies should also be performed at End of Treatment; if imaging studies were performed within 4 weeks prior to the end of treatment, a repeat evaluation does not need to be performed at this time, unless clinically indicated. Whenever clinically feasible, it is strongly requested that the sameimaging modality (MRI or CT) at both baseline and subsequent timepoints be utilized for any given patient throughout their evaluation on-study.

11 Peripheral blood (serum) samples (10 mL) will be collected for the detection of tagraxofusp reactive antibodies at the completion of every cycle and at specified pointsthereafter.

12 Patients will be followed for survival approximately every 90 days, until assessment of the primary objective for the study is complete for all treated patients; please consult the appropriate protocol section for description of situations in which survival and other data may be collected following study completion. Patients who undergo SCT will befollowed for the occurrence of veno-occlusive disease as part of long-term follow-up. A blood sample for immunogenicity studies will be collected at least 16 weeks to up to 20 weeks after the last tagraxofusp dose.

13 Premedications administered 60 minutes (±15 minutes) before tagraxofusp treatment: Diphenhydramine 50 mg IV (or an equivalent dose of another H_1-histamine antagonist);acetaminophen 650 mg orally (or equivalent dose of paracetamol); methylprednisolone 50 mg IV (or an equivalent dose of another corticosteroid); famotidine 20 mg IV (or an equivalent doses of another H_2-histamine antagonist).

14 Following treatment with premedication, tagraxofusp will be administered as a 15-minute infusion once daily for the first 3 consecutive days of a 28-day cycle. Individualtagraxofusp infusions may be delayed to allow for toxicity resolution, but all 3 infusions should be completed within 10 days.

15 Patients with MF should complete the MPN-SAF TSS (10-question evaluation concerning symptoms/well-being during the prior week) with every treatment cycle, prior toreceiving therapy on the subsequent cycle. The MPN-SAF TSS should also be completed at End of Treatment.

16 Tumor response assessments are to be conducted every 12 weeks (±7 days) and at End of Treatment, detailed in protocol Section 11.12 and Section 18.1.

Table 7-3 Stage 3A: Study Events Schedule for Screening and Cycles 1-4 for Patients with CMML

Procedures	Day -28 to 0 *Screening*	Days 1-3 (Up to Day 10 if Infusion(s) Held) Tagraxofusp Treatment *Pre-Infusion*	*Tagraxofusp Infusion*	Cycle 1 Day 8±3[1]	Day 21±3[2] *End of Cycle*[2]	Day 28±3, Then Every 7±3 days *(Only if Delayed End of Cycle: for Toxicity Resolution)*
Informed consent form[3]	X					
Inclusion/exclusion criteria[3]	X					
Medical history	X					
12-lead ECG	X					
ECOG performance status	X	X (Infusion 1, C1)			X	X
Physical examination, including assessment for hepatomegaly and splenomegaly	X			X	X	
Pregnancy test[4]	X					
Vital signs and weight[5]	X	X	X	X	X	X
MUGA scan or 2-D Echocardiogram[6]	X					
Hematology[7]	X	X		X	X	X
Serum chemistry[8]	X	X		X	X	X
Coagulation parameters: PT/INR, aPTT[9]	X	X		X	X	X
Peripheral blood for flow cytometry	X	X (Infusion 1)		X		
Peripheral blood for IL-10 testing		X (Infusion 1, C1)			X C1 and C4	
Peripheral blood for translational studies (select centers)		X (Infusion 1, C1)			X C1 and C4	
Bone marrow aspiration/biopsy[10]	X				X C1 and C4[11]	
MRI or CT for assessment of hepatomegaly and splenomegaly [12]	X				X C1 and C4[11]	
Immunogenicity sampling[13]		X (Infusion 1)			X	

Procedures	Day -28 to 0	Days 1-3 (Up to Day 10 if Infusion(s) Held) Tagraxofusp Treatment		Cycle 1 Day 8±3[1]	Day 21±3[2]	Day 28±3, Then Every 7±3 days *(Only if Delayed End of Cycle: for Toxicity Resolution)*
	Screening	*Pre-Infusion*	*Tagraxofusp Infusion*		*End of Cycle*[2]	
Cytogenetic and molecular genetic testing	X				X C1 and C4	
Administration of premedications[14]		X				
Tagraxofusp administration[15]			X			
MPN-SAF TSS evaluation[16]	X				X	
Baseline/tumor response assessment[17]	X				X C1 and C4[11]	
Vision assessment	X				X	
Prior/concomitant medications/therapies (including concomitant platelet and RBC transfusions)	X	X	X	X	X	X
AE and SAE monitoring	X	X	X	X	X	X

AE = adverse event; aPTT = activated partial thromboplastin time; CT = computed tomography; ECG = electrocardiogram; ECOG = Eastern Cooperative Oncology Group; INR = international normalized ratio; MPN-SAF TSS = Myeloproliferative Neoplasm Symptom Assessment Form Total Symptom Score; MUGA = multigated acquisition; PT = prothrombin time; SAE = serious adverse event.

1 Day 8 visit is required for C1 only and must be performed at the site.

2 The end-of-cycle evaluations (Day 21 or thereafter) may also serve as the pre-infusion evaluations for the subsequent cycles; with the exception of vital signs, these assessments do not need to be duplicated on successive days unless there is an abnormality or other clinically relevant reason for repeat evaluation.

3 Refer to protocol Section 16.3 (Patient Information and Informed Consent) and Section 7 (Populations to be Studied) for details.

4 Urine or serum pregnancy test must be performed within 1 week prior to treatment for WOCBP.

5 Height will be measured at screening only; weight does not need to be measured more than once per day, and should be measured pre-infusion on treatment days. Vital signs should be performed after patient is sitting for 3 to 5 minutes. During dosing period, vital signs should be taken immediately prior to infusion, at 0 (i.e., immediately after completion of infusion), and at 30 and 60 minutes post-infusion.

6 A MUGA scan or 2-D ECHO to quantify LVEF must be completed within 28 days prior to start of first cycle of study drug.

7 To be collected prior to tagraxofusp infusion if during dosing period. Hematology includes WBC count with differential, RBC count, hematocrit, Hb, platelet count, and immature myeloid cells (including blasts + myelocytes +metamyelocytes + promyelocytes + nucleated red blood cells) (for safety assessments as well as efficacy assessments at the end of each cycle in between bone marrow assessments).

8 To be collected prior to tagraxofusp infusion if during dosing period. Serum chemistry incudes electrolytes and additional parameters (equivalent to Chem-20): ALT, albumin, ALP, AST, bicarbonate, bilirubin, BUN, calcium, magnesium, chloride, creatinine, glucose, LDH, phosphate, potassium, sodium, total protein, uric acid, and CPK. See protocol Appendix Section 18.3 for administration of albumin if serum albumin decreases to < 3.0 g/dL (< 30 g/L) during treatment days or in the immediate post-treatment period.

9 In lieu of PT, INR may be measured.

10 Morphology and WBC differential/blast count on aspirate. Baseline must be performed within 28 days prior to the first administration of tagraxofusp. Subsequent bone marrow aspirates and biopsies will be performed at the end of C1 and C4 and then every 12 weeks (±7 days) thereafter. Bone marrow evaluation should also be performed at End of Treatment including patients who discontinue study therapy prior to completion of C4; if a bone marrow evaluation was performed within 4 weeks prior to the end of treatment, a repeat evaluation does not need to be performed at this time, unless clinically indicated.

11 Bone marrow aspirate and imaging studies are to be repeated at the end of C1 and C4 and then every 12 weeks (±7 days) thereafter.

12 A baseline MRI or CT scan must be performed within 28 days prior to the first administration of tagraxofusp. Subsequent imaging studies will be performed at the end of C1 and C4 and then every 12 weeks (±7 days) until there is evidence of relapsed or progressive disease. Abdominal scan is required; evaluation of chest/pelvis may be obtained at the Investigator's discretion. Scan must include measurement of spleen and liver volume. Subsequent to Week 24, for patients receiving ongoing therapy (or those who discontinue therapy but are without evidence of PD and continue to be evaluated in ongoing follow-up) the frequency of imaging studies may be reduced (obtained less frequently) if peripheral blood and/or clinical assessments indicate clinical stability, as per the Investigator's discretion, in consultation with the Medical Monitor. Imaging should also be performed at End of Treatment, including patients who discontinue study therapy prior to completion of C4; if a imaging was performed within 4 weeks prior to the end of treatment, a repeat evaluation does not need to be performed at this time, unless clinically indicated. Whenever clinically feasible, it is strongly requested that the same imaging modality (MRI or CT) at both baseline and subsequent timepoints be utilized for any given patient throughout their evaluation on-study.

13 Peripheral blood (serum) samples (10 mL) will be collected for the detection of tagraxofusp reactive antibodies on Day 1 (pre-infusion) and Day 21 (the end of cycle collection may also serve as the pre-infusion collection for the subsequent cycle).

14 Premedications administered 60 minutes (±15 minutes) before tagraxofusp treatment: Diphenhydramine 50 mg IV (or an equivalent dose of another H_1-histamine antagonist); acetaminophen 650 mg orally (or equivalent dose of paracetamol); methylprednisolone 50 mg IV (or an equivalent dose of another corticosteroid); famotidine 20 mg IV (or an equivalent dose of another H_2-histamine antagonist).

15 Following treatment with premedication, tagraxofusp will be administered as a 15-minute infusion once daily for the first 3 consecutive days of a 21-day cycle. Individual tagraxofusp infusions may be delayed to allow for toxicity resolution, but all 3 infusions should be completed within 10 days.

16 Patients complete the MPN-SAF TSS (10-question evaluation concerning symptoms/well-being during the prior week) during screening and with every treatment cycle, prior to receiving therapy on the subsequent cycle. The MPN-SAF TSS should also be completed at End of Treatment.

17 Tumor assessments will be performed during screening, at the end of C1 and C4, then every 12 weeks (±7 days) and at End-of-Treatment per the 2015 MDS/MPN criteria, as detailed in protocol Section 11.12 and Section 18.1

Table 7-4 Stage 3A: Study Events Scheduled for Cycle 5 and Beyond for Patients with CMML

Procedures	Cycle 5 and Beyond			Cycles 5 and Beyond[2] 12 Weeks (i.e., every 3 cycles) ±7 Days	End-of-treatment	Safety: Through 30 Days After Last Infusion	Survival: Every 90 Days After Last Infusion
	Days 1-3 (Up to Day 10 if Infusion(s) Held) Tagraxofusp Treatment		Day 28±3[1]				
	Pre-Infusion	*Infusion*	*End of Cycle*[1]				
Concomitant medications/therapies (including platelet and RBC transfusions)	X		X		X		
AE and SAE monitoring	X	X	X	X	X	X	
ECOG performance status			X		X		
Physical examination including assessment of hepatomegaly and splenomegaly			X		X		
Vital signs and weight[3]	X	X	X		X		
Hematology[4]	X		X		X		
Serum chemistry[5]	X		X		X		
Coagulation parameters: PT/INR, aPTT[6]	X		X		X		
Peripheral blood for flow cytometry	X (Infusion 1)				X		
Bone marrow aspiration/biopsy[7]				X[8]	X		
MRI or CT for assessment of hepatomegaly and splenomegaly[9]				X[8]	X		
Immunogenicity sampling[10]			X		X	X	X[11]
Cytogenetic and molecular genetic testing				X	X		
Administration of premedications[12]	X						
Administration of Tagraxofusp[13]		X					
MPN-SAF TSS evaluation[14]			X		X		
Vision assessment			X		X		
Tumor response assessment[15]				X[8]	X		
Long-term follow-up[11]						X	X

AE = adverse event; aPTT = activated partial thromboplastin time; CT = computed tomography; ECG = electrocardiogram; ECOG = Eastern Cooperative Oncology Group;

INR = international normalized ratio; MPN-SAF TSS = Myeloproliferative Neoplasm Symptom Assessment Form Total Symptom Score; PT = prothrombin time; SAE = serious adverse event.

1 The end-of-cycle evaluations (Day 28 or thereafter) may also serve as the pre-infusion evaluations for the subsequent cycles; with the exception of vital signs, these assessments do not need to be duplicated on successive days unless there is an abnormality or other clinically relevant reason for repeat evaluation.

2 Disease evaluation via bone marrow aspirate/biopsy and/or imaging studies (when relevant) will occur every 12 weeks (±7 days). The frequency of bone marrow evaluation may be reduced (obtained less frequently) if peripheral blood, clinical and/or imaging assessments indicate clinical stability, as per the Investigator's discretion in consultation with the Medical Monitor. The frequency of imaging assessments may also be reduced if clinical/laboratory findings indicate stability, as per the Investigator's discretion in consultation with the Medical Monitor.

3 Height will be measured at screening only; weight does not need to be measured more than once per day, and should be measured pre-infusion on treatment days. Vital signs should be performed after patient is sitting for 3 to 5 minutes. If during dosing period, vital signs should be taken immediately prior to infusion, at 0 (i.e., immediately after completion of infusion), and at 30 and 60 minutes post-infusion.

4 To be collected prior to tagraxofusp infusion if during dosing period. Hematology includes WBC count with differential, RBC count, hematocrit, hemoglobin, platelet count, and immature myeloid cells (including blasts + myelocytes +metamyelocytes + promyelocytes + nucleated red blood cells) (for safety assessments as well as efficacy assessments at the end of each cycle in between bone marrow assessments).

5 To be collected prior to tagraxofusp infusion if during dosing period. Serum chemistry incudes electrolytes and additional parameters (equivalent to Chem-20): ALT, albumin, alkaline phosphatase, AST, bicarbonate, bilirubin, BUN, calcium, magnesium, chloride, creatinine, glucose, LDH, phosphate, potassium, sodium, total protein, uric acid, chloride, and CPK. See protocol Appendix Section 18.3 for administration of albumin if serum albumin decreases to < 3.0 g/dL (< 30 g/L) during treatment days or in the immediate post-treatment period.

6 In lieu of PT, INR may be measured.

7 Morphology and WBC differential /blast count on aspirate. In general, bone marrow aspirates and biopsies will be performed at the end of C1 and C4 and then every 12 weeks (±7 days) thereafter. Bone marrow evaluation should also be performed at End of Treatment; if a bone marrow evaluation was performed within 4 weeks prior to the end of treatment, a repeat evaluation does not need to be performed at this time, unless clinically indicated.

8 Bone marrow aspirate and imaging studies are to be repeated at the end of C1 and C4 and then every 12 weeks (±7 days) thereafter.

9 Imaging studies will be performed at the end of C1 and C4 and then every 12 weeks (±7 days) until there is evidence of relapsed or progressive disease. Abdominal scan is required; evaluation of chest/pelvis may be obtained at the Investigator's discretion. Scan must include measurement of spleen and liver volume. Imaging studies should also be performed at End of Treatment; if imaging was performed within 4 weeks prior to the end of treatment, a repeat evaluation does not need to be performed at this time, unless clinically indicated. Whenever clinically feasible, it is strongly requested that the same imaging modality (MRI or CT) at both baseline and subsequent timepoints be utilized for any given patient throughout their evaluation on-study.

10 Peripheral blood (serum) samples (10 mL) will be collected for the detection of tagraxofusp reactive antibodies at the completion of every cycle and at specified points thereafter.

11 Patients will be followed for survival approximately every 90 days, until assessment of the primary objective for the study is complete for all treated patients; please consult the appropriate protocol section for description of situations in which survival and other data may be collected following study completion. Patients who undergo SCT will be followed for the occurrence of veno-occlusive disease as part of long-term follow-up. A blood sample for immunogenicity studies will be collected at least 16 weeks to up to 20 weeks after the last tagraxofusp dose.

12 Premedications administered 60 minutes (±15 minutes) before tagraxofusp treatment: Diphenhydramine 50 mg IV (or an equivalent dose of another H1-histamine antagonist); acetaminophen 650 mg orally (or equivalent dose of paracetamol); methylprednisolone 50 mg IV (or an equivalent dose of another corticosteroid); famotidine 20 mg IV (or an equivalent doses of another H2-histamine antagonist).

13 Following treatment with premedication, tagraxofusp will be administered as a 15-minute infusion once daily for the first 3 consecutive days of a 28-day cycle. Individual

tagraxofusp infusions may be delayed to allow for toxicity resolution, but all 3 infusions should be completed within 10 days.

14 Patients should complete the MPN-SAF TSS (10-question evaluation concerning symptoms/well-being during the prior week) with every treatment cycle, prior to receiving therapy on the subsequent cycle. The MPN-SAF TSS should also be completed at End of Treatment.

15 Tumor response assessments are to be conducted every 12 weeks (±7 days) and at End of Treatment, detailed in protocol Section 11.12, and Section 18.1.

8. RESPONSE CRITERIA

8.1. Myelofibrosis

Table 8-1 Stage 2: Revised IWG-MRT and ELN Response Criteria for MF

Response categories	Required criteria (for all response categories, benefit must last for ≥12 weeks to qualify as a response)
CR	Bone marrow [1]: Age-adjusted normocellularity; < 5% blasts; ≤ Grade 1 MF [2] and Peripheral blood: Hemoglobin ≥ 100 g/L and < ULN; neutrophil count ≥ $1x10^9$/L and < ULN; Platelet count ≥ $100x10^9$/L and < ULN; < 2% immature myeloid cells[3] and Clinical: Resolution of disease symptoms; spleen and liver not palpable; no evidence of EMH
PR	Peripheral blood: Hemoglobin ≥100 g/L and < ULN; neutrophil count ≥$1x10^9$/L and < ULN; Platelet count ≥$100x10^9$/L and < ULN; < 2% immature myeloid cells[3] and Clinical: Resolution of disease symptoms; spleen and liver not palpable; no evidence of EMH or Bone marrow[1]: Age-adjusted normocellularity; < 5% blasts; ≤ Grade 1 MF[2] and Peripheral blood: Hemoglobin ≥85g/L but < 100g/L and < ULN; neutrophil count ≥$1×10^9$/L and < ULN; platelet count ≥$50×10^9$/L but < $100×10^9$/L and < ULN; < 2% immature myeloid cells[3] and Clinical: Resolution of disease symptoms; spleen and liver not palpable; no evidence of EMH
Clinical improvement (CI)	The achievement of anemia, spleen or symptom response without progressive disease or increase in severity of anemia, thrombocytopenia or neutropenia[4]
Anemia response	Transfusion independent patients: a ≥ 20g/L increase in hemoglobin level[5] Transfusion dependent patients: becoming transfusion independent[6]
Spleen response[7]	A baseline splenomegaly that is palpable 5-10cm below the LCM, becomes not palpable[8] or A baseline splenomegaly that is palpable > 10cm below the LCM, decreases by ≥ 50%[8] A baseline splenomegaly that is palpable < 5cm below the LCM is not eligible for spleen response A spleen response requires confirmation by MRI or CT showing ≥ 35% spleen volume reduction
Symptom response	A ≥ 50% reduction in the MPN-SAF TSS[9]

Response categories	Required criteria (for all response categories, benefit must last for ≥12 weeks to qualify as a response)
Progressive disease[10]	Appearance of new splenomegaly that is palpable at least 5cm below the LCM or A ≥ 100% increase in palpable distance, below LCM, for baseline splenomegaly of 5-10 cm or A 50% increase in palpable distance, below LCM, for baseline splenomegaly of > 10 cm or Leukemic transformation confirmed by a bone marrow blast count of ≥ 20% or A peripheral blast content of ≥ 20% associated with an absolute blast count of ≥1x109/L that lasts for at least 2 weeks
Stable disease	Belonging to none of the above listed response categories
Relapse	No longer meeting criteria for at least CI after achieving CR, PR or CI, or Loss of anemia response persisting for at least 1 month or Loss of spleen response persisting for at least 1 month
Cytogenetic remission	At least 10 metaphases must be analyzed for cytogenetic response evaluation and requires confirmation by repeat testing within a 6-month window CR: eradication of a preexisting abnormality PR: ≥ 50% reduction in abnormal metaphases (partial response applies only to patients with at least 10 abnormal metaphases at baseline)
Molecular remission	Molecular response evaluation must be analyzed in peripheral blood granulocytes and requires confirmation by repeat testing within a 6-month window CR: eradication of a preexisting abnormality PR: ≥50% decrease in allele burden (partial response applies only to patients with at least 20% mutant allele burden at baseline)
Cytogenetic/molecular relapse	Re-emergence of a pre-existing cytogenetic or molecular abnormality that is confirmed by repeat testing

Adapted from Tefferi, et al. Blood 2013; 122(8):1395-98.
EMH: extramedullary hematopoiesis (no evidence of EMH implies the absence of pathology- or imaging study-proven nonhematosplenic EMH). LCH: left costal margin. ULN: upper limit of normal.

1 Baseline and posttreatment bone marrow slides are to be interpreted at one sitting by a central review process. Cytogenetic and molecular responses are not required for CR assignment.

2 Grading of MF is according to the European classification. Thiele et al. European consensus on grading bone marrow fibrosis and assessment of cellularity. Haematologica 2005; 90: 1128.
It is underscored that the consensus definition of a CR bone marrow is to be used only in those patients in which all other criteria are met, including resolution of leukoerythroblastosis. It should also be noted that it was a particularly difficult task for the working group to reach a consensus regarding what represents a complete histologic remission.

3 Immature myeloid cells constitute blasts + promyelocytes + myelocytes + metamyelocytes + nucleated red blood cells. In splenectomized patients, < 5% immature myeloid cells is allowed.
4 See above for definitions of anemia response, spleen response, and progressive disease. Increase in the severity of anemia constitutes the occurrence of new transfusion dependency or a >= 20 g/L decrease in hemoglobin level from pretreatment baseline that lasts for at least 12 weeks. Increase in severity of thrombocytopenia or neutropenia is defined as a 2-grade decline, from pretreatment baseline, in platelet count or absolute neutrophil count, according to the CTCAE version 4.03. In addition, assignment to CI requires a minimum platelet count of >= 25×10^9/L and absolute neutrophil count of >=$0.5x10^9$/L.
5 Applicable only to patients with baseline hemoglobin of < 100g/L. In patients not meeting the strict criteria for transfusion dependency at the time of study enrollment (see as follows), but have received transfusions within the previous month, the pretransfusion hemoglobin level should be used as the baseline.
6 Transfusion dependency before study enrollment is defined as transfusions of at least 6 units of pRBCs in the 12 weeks prior to study enrollment, for a hemoglobin level of < 85 g/L, in the absence of bleeding or treatment-induced anemia. In addition, the most recent transfusion episode must have occurred in the 28 days prior to study enrollment. Response in transfusion-dependent patients requires absence of any pRBC transfusions during any consecutive "rolling" 12-week interval during the treatment phase, capped by a hemoglobin level of >=85g/L.
7 In splenectomized patients, palpable hepatomegaly is substituted with the same measurement strategy.
8 Spleen or liver responses must be confirmed by imaging studies where a >=35% reduction in spleen volume, as assessed by MRI or CT, is required. Furthermore, a >=35% volume reduction in the spleen or liver, by MRI or CT, constitutes a responseregardless of what is reported with physical examination.
9 Symptoms are evaluated by the MPN-SAF TSS. The MPN-SAF TSS is assessed by the patients themselves and this includes fatigue, concentration, early satiety, inactivity, night sweats, itching, bone pain, abdominal discomfort, weight loss and fevers. Scoring is from 0 (absent/as good as it can be) to 10 (worst imaginable/bad as it can be) for each item. The MPN-SAF TSS is the summation of the individual scores (0-100 scale). Symptom response requires >=50% reduction in theMPN-SAF TSS.
10 Progressive disease assignment for splenomegaly requires confirmation by MRI or CT showing a >=25% increase in spleen volume from baseline. Baseline values for both physical examination and imaging studies refer to pretreatment baseline and not to posttreatment measurements.

Table 8-2 Grading of Myelofibrosis

MF-0	Scattered linear reticulin with no intersections (crossovers) corresponding to normal BM.
MF-1	Loose network of reticulin with many intersections, especially in perivascular areas.
MF-2	Diffuse and dense increase in reticulin with extensive intersections, occasionally with focal bundles of thick fibers mostly consistent with collagen, and/or focal osteosclerosis.[1]
MF-3	Diffuse and dense increase in reticulin with extensive intersections and coarse bundles of thick fibers consistent with collagen, usually associated with osteosclerosis.[1]

Adapted from Arber et al, Blood 2016; 127(20): 2391-2405

Semiquantitative grading of BM fibrosis (MF) with minor modifications concerning collagen and osteosclerosis. Fiber densityshould be assessed only in hematopoietic areas.

1. In grades MF-2 or MF-3 an additional trichrome stain is recommended.

8.2. Chronic Myelomonocytic Leukemia

Table 8-3 Stage 2: Modified IWG Response Criteria for Altering Natural History of MDS (Including CMML)

Category	Response criteria (response must last at least 4 weeks)
Complete remission	Bone marrow: ≤5% myeloblasts with normal maturation of all cell lines(1) Persistent dysplasia will be noted[1,2] Peripheral blood[3]: Hemoglobin ≥11g/dL Platelets ≥100×10^9/L Neutrophils ≥1.0×10^9/L[2] Blasts 0%
Partial remission	All CR criteria if abnormal before treatment except: Bone marrow blasts decreased by ≥50% but still > 5% Cellularity and morphology not relevant
Bone Marrow CR(2)	Bone marrow ≤5% myeloblasts and decrease by ≥50% over pretreatment[2] Peripheral blood: If hematologic improvement responses, they will be noted in addition to marrow CR[2]
Stable disease	Failure to achieve at least PR, but no evidence of progression for > 8 weeks
Failure	Death during treatment or disease progression characterized by worsening of cytopenias, increase in percentage of bone marrow blasts or progression to a more advanced MDS FAB subtype than pretreatment
Relapse after CR or PR	At least 1 of the following: Return to pretreatment bone marrow blast percentage Decrement of ≥50% from maximum remission/response levels in granulocytes or platelets Reduction in Hemoglobin concentration by ≥1.5 g/dL or transfusion dependence
Cytogenetic response	Complete: Disappearance of the chromosomal abnormality without appearance of new ones Partial: At least 50% reduction of the chromosomal abnormality
Disease progression	For patients with: Less than 5% blasts: ≥ 50% increase in blasts to > 5% blasts 5-10% blasts: ≥ 50% increase to > 10% blasts 10-20% blasts: ≥ 50% increase to > 20% blasts 20-30% blasts: ≥ 50% increase to > 30% blasts Any of the following: At least 50% decrement from maximum remission/response in granulocytes or platelets Reduction in Hemoglobin by ≥ 2 g/dL Transfusion dependence

Survival	Endpoints: Overall: death from any cause Event free: failure or death from any cause PFS: disease progression or death from MDS DFS: time to relapse Cause-specific death: death related to MDS

Adapted from Cheson, et al. Blood 2006; 108(2):419-25.
Deletions to the IWG response criteria are not shown. To convert hemoglobin from g/dL to g/L, multiply g/dL by 10.
MDS: myelodysplastic syndromes. Hgb: hemoglobin. CR: complete remission. HI: Hematologic improvement. PR: partial remission. FAB: French-American-British. AML: Acute myeloid leukemia. PFS: Progression-free survival. DFS: Disease-free survival.

1 Dysplastic changes should consider the normal range of dysplastic changes (modification) (Ramos 1999)
2 Modification to IWG response criteria
3 In some circumstances, protocol therapy may require the initiation of further treatment (e.g., consolidation, maintenance) before the 4-week period. Such patients can be included in the response category into which they fit at the time the therapy is started. Transient cytopenias during repeated chemotherapy courses should not be considered as interrupting durability of response, as long as they recover to the improved counts of the previous course.

Table 8-4 Stage 3A: 2015 MDS/MPN Response Criteria (Savona 2015)

Criteria for measurement of treatment response in adult MDS/MPN

CR (presence of all of the following improvements)[1]

- Bone marrow:
  - ≤ 5% myeloblasts (including monocytic blast equivalent in case of CMML)
  - Normal maturation of all cell lines and return to normal cellularity[1]
  - Osteomyelofibrosis absent or equal to "mild reticulin fibrosis" (≤ grade 1 fibrosis)[2]
- Peripheral blood[3]
  - WBC ≤ 10×10^9 cells/L
  - Hgb ≥ 11 g/dL
  - Platelets ≥ 100×10^9/L; ≤ 450×10^9/L
  - Neutrophils ≥ 1.0×10^9/L
  - Blasts 0%
  - Neutrophil precursors reduced to ≤ 2%
  - Monocytes ≤ 1×10^9/L
- Extramedullary disease: Complete resolution of extramedullary disease present before therapy (e.g. cutaneous disease, disease-related serous effusions), including palpable hepatosplenomegaly
- Persistent low-level dysplasia is permitted given subjectivity of assignment of dysplasia[1]

Note: Provisional category of CR with resolution of symptoms is removed from the original criteria.[3]

Complete cytogenetic remission

- Resolution of previously present chromosomal abnormality (known to be associated with myelodysplastic, syndrome myeloproliferative neoplasms, or MDS/MPN), as seen on classic karyotyping with minimal of 20 metaphases or FISH[4]

Criteria for measurement of treatment response in adult MDS/MPN

Partial remission

- Normalization of peripheral counts and hepatosplenomegaly with bone marrow blasts (and blast equivalents) reduced by 50%, but remaining > 5% of cellularity except in cases of MDS/MPN with ≤ 5% bone marrow blasts at baseline

Marrow response

- Optimal marrow response: Presence of all marrow criteria necessary for CR without normalization of peripheral blood indices as presented above
- Partial marrow response: Bone marrow blasts (and blast equivalents) reduced by 50%, but remaining > 5% of cellularity, or reduction in grading of reticulin fibrosis from baseline on at least 2 bone marrow evaluations spaced at least 2 months apart

Clinical benefit

Requires 1 of the following in the absence of progression or CR/partial response and independent of marrow response (response must be verified at ≥ 8 weeks) to be considered a clinical benefit. (Note: "Cord blood response" was corrected to "response".)

- Erythroid response
  - Hgb increase by ≥ 2.0 g/dL
  - Transfusion independence for ≥ 8 weeks for patients requiring at least 4 packed red blood cell transfusions in the previous 8 weeks
  - Only red blood cell transfusions given based on physician's judgement for a pretreatment Hgb of ≤ 8.5 g/dL will count in the red blood cell TI response evaluation[5]
- Platelet response
  - Transfusion independence when previously requiring platelet transfusions of at least a rate of 4 platelet transfusions in the previous 8 weeks
  - Pretreatment ≤ 20×10^9/L: increase from < 20×10^9/L to > 20×10^9/L and by at least 100%
  - Pretreatment > 20×10^9/L but ≤ 100×10^9/L: absolute increase of ≥ 30×10^9/L[5]
- Neutrophil response
  - Pretreatment ≤ 0.5×10^9/L, at least 100% increase and an absolute increase ≥ 0.5×10^9/L
  - Pretreatment, > 0.5×10^9/L and ≤ 1.0×10^9/L, at least 50% increase and an absolute increase ≥ 0.5×10^9/L[5]
- Spleen response
  - Either a minimum 50% reduction in palpable splenomegaly of a spleen that is at least 10 cm at baseline or a spleen that is palpable at more than 5 cm at baseline becomes not palpable

Symptom response (improvement in symptoms as noted by decrease of ≥ 50% as per the MPN-SAF TSSscoring < 20 were not considered eligible for measuring clinical benefit)[6]

1 Presence of dysplastic changes, which may be interpreted within the scope of normal range of dysplastic changes, may still exist in the presence of CR as allowed in MDS IWG. Marrow should exhibit age-adjusted normocellularity in CR.

2 If there is no significant fibrosis present on the initial bone marrow biopsy, a second biopsy is not required to prove resolution of fibrosis. Grading of fibrosis in measurement of treatment response should be according to the European Consensus System (Thiele 2005).

3 Given the current lack of a validated tool to assess complete resolution of symptoms in MDS/MPN, "CR with resolution of symptoms" (a complete resolution of disease-related symptoms as noted by the MPN-SAF TSS in presence of CR) will be a

provisional category of disease response.

4 Loss of cytogenetic burden of disease (via FISH or classic karyotyping) known to adversely affect prognosis is required to reach complete cytogenetic remission. Decrease in the cytogenetic burden of disease must be by ≥ 50% (via FISH or classic karyotyping) to be indicative of a partial cytogenetic response. Given variability of fluorescent probes used in FISH, cytogenetic normalization via FISH will depend on the performance characteristics of the specific probes used.

5 Resolution of abnormal peripheral blood counts must persist for at least 2 separate analyses over at least 8 weeks. In the case of proliferative MDS/MPN, CR will include resolution of thrombocytosis to a normal platelet count (150-450×10^9/L) and resolution of leukocytosis to WBC ≤10×10^9 cells/L but ≥1.5×10^9/L. Hemoglobin should be maintained >11 g/dL and platelets ≥100×10^9/L without the support of transfusions. Clinical benefit may occur when these changes occur in absence of other changes required for CR or marrow response. Platelet and packed red blood cell TI would be considered for clinical benefit, and duration of TI should be monitored. Reduction in myeloid precursors (promyelocytes, myelocytes, metamyelocytes, nucleated red blood cells) to less than appreciable levels (≤ 2-3%) and/or 1 x 10^9/L monocytosis in the absence of infection, cytokine treatment, or other reactive causes.

6 MPN-SAF TSS validation among patients with MDS/MPN is currently under way (R.A. Mesa, personal communication, 2014).

Criteria for measurement of disease progression in adult MDS/MPN

Combination of 2 major criteria, 1 major and 2 minor criteria, or 3 minor criteria from list

Major criteria

- Increase in blast count[1]
  - < 5% blasts: ≥ 50% increase and to > 5% blasts
  - 5-10% blasts: ≥ 50% increase and to > 10% blasts
  - 10-20% blasts: ≥ 50% increase and to > 20% blasts
  - 20-30% blasts: ≥ 50% increase and to > 30% blasts [2]
- Evidence of cytogenetic evolution[3]
  - Appearance of a previously present or new cytogenetic abnormality in complete cytogenetic remission via FISH or classic karyotyping.
  - Increase in cytogenetic burden of disease by ≥ 50% in partial cytogenetic remission via FISH or classic karyotyping.
- New extramedullary disease.
  - Worsening splenomegaly:

    Progressive splenomegaly that is defined by IWG-MRT: the appearance of a previously absent splenomegaly that is palpable at > 5 cm below the left costal margin or a minimum 100% increase in palpable distance for baseline splenomegaly of 5-10 cm or a minimum 50% increase in palpable distance for baseline splenomegaly of > 10 cm

    Extramedullary disease outside of the spleen to include new/worsening hepatomegaly, granulocytic sarcoma, skin lesions, etc.

Minor criteria

- Transfusion dependence[4]
- Significant loss of maximal response on cytopenias ≥ 50% decrement from maximum remission/response in granulocytes or platelets
- Reduction in Hgb by ≥ 1.5 g/dL from best response or from baseline as noted on complete blood count
- Increasing symptoms as noted by increase in ≥ 50% as per the MPN-SAF TSS[5]
- Evidence of clonal evolution (molecular)[6]

1 Blasts as measured from the bone marrow.

2 Patients with development of acute myeloid leukemia from MDS/MPN; 20-30% blasts may be allowed on some clinical trials for patients with MDS/MPN.

3 Increase in cytogenetic burden of disease by ≥ 50% (via FISH or classic karyotyping). Given variability of fluorescent probes used in FISH, cytogenetic normalization via FISH will depend on specific probes used.

4 Transfusion dependency is defined by a history of at least 2 units of red blood cell transfusions in the past month for a hemoglobin level < 8.5 g/dL that was not associated with clinically overt bleeding. Cytopenia resulting from therapy should not be considered in assessment of progression.

5 MPN-SAF TSS validation among patients with MDS/MPN is currently under way (R.A. Mesa, personal communication, 2014).

6 The identification of new abnormalities using single nucleotide polymorphism arrays or sequencing or a clearly significant increase in mutational burden of a previously detected abnormality. Precise criteria for defining new abnormalities and what exactly constitutes a significant increase in mutational burden are open to interpretation; we suggest that this criterion should be used conservatively based on current evidence.

9. REVISION HISTORY

STML-401-0314 SAP was revised to Version 2.0 on 30 November 2023. The following analyses were added:

- New column structures were specified (Section 3.5)
- Additional analyses of baseline or disease history data were added (Section 4.2)
- Time to event and additional duration of response parameters were specified (Sections 4.3.1 and 4.3.2)
- Analyses of Spleen volume were clarified (Section 4.3.5)
- Analyses of Tumor Symptom Score were clarified (Section 4.3.6)
- More details on transfusion dependency and rate calculation were added (Section 4.3.7)
- Laboratory results related to efficacy were added (Section 4.3.8)
- Additional adverse event analyses were added (Section 4.4.2)
- Analyses of laboratory results were simplified (Section 4.4.3)
- Analysis of pre-treatment medications was removed and analyses of prior systemic therapy were added (Section 4.4.6)
- Some immunogenicity analyses were removed (Section 4.5)
- Methods for pharmacokinetic NCA were specified (Section 1.2.5.2). Summary of pharmacokinetic analyses was added (Section 4.6)
- Statistical Table, Listing, and Figure shells are moved to a separate document from Section 10